

# Best Available Copy

(19)  Canadian Intellectual Property Office  
An Agency of Industry Canada

Office de la Propriété Intellectuelle du Canada  
Un organisme d'Industrie Canada

(11) CA 2 303 526 (13) A1

(40) 03.10.2000  
(43) 03.10.2000

(12)

(21) 2 303 526

(51) Int. Cl.?

C12N 7/02, C12N 5/10,  
C12N 15/51

(22) 31.03.2000

(30) 199 15 178.4 DE 03.04.1999

(72)

BARTENSCHLAGER, RALF (DE).

(71) BARTENSCHLAGER, RALF,  
Nach dem Alten Schloss 22  
D-55239, GAU-ODERNHEIM, XX (DE).

(74)

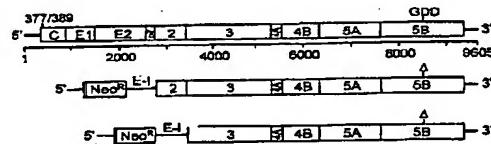
BORDEN LADNER GERVAIS LLP

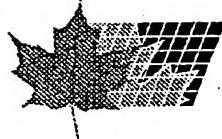
(54) SYSTEME DE CULTURE DU VIRUS DE L'HEPATITE C

(54) HEPATITIS C VIRUS CELL CULTURE SYSTEM

(57)

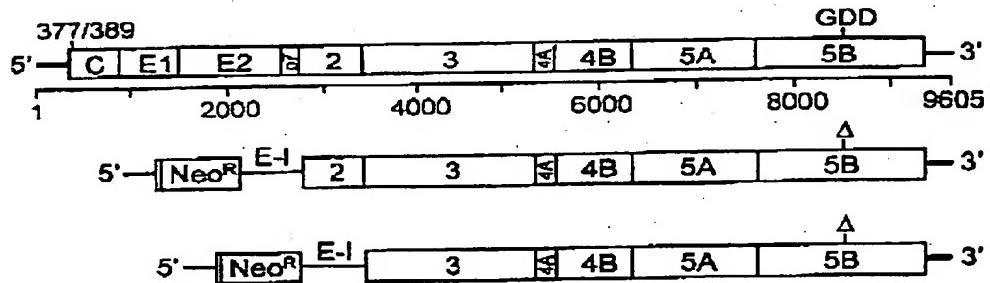
The hepatitis C virus (HCV) cell culture system according to the invention consists of human hepatoma cells, which are transfected with a HCV-RNA construct, that comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR as well as a minimum of one marker gene for selection (selection gene).





(21)(A1) 2,303,526  
(22) 2000/03/31  
(43) 2000/10/03

(72) BARTENSCHLAGER, RALF, DE  
(71) BARTENSCHLAGER, RALF, DE  
(51) Int.Cl.<sup>7</sup> C12N 7/02, C12N 15/51, C12N 5/10  
(30) 1999/04/03 (199 15 178.4) DE  
(54) SYSTEME DE CULTURE DU VIRUS DE L'HEPATITE C  
(54) HEPATITIS C VIRUS CELL CULTURE SYSTEM



(57) The hepatitis C virus (HCV) cell culture system according to the invention consists of human hepatoma cells, which are transfected with a HCV-RNA construct, that comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR as well as a minimum of one marker gene for selection (selection gene).



**Hepatitis C Virus Cell Culture System**

5

**ABSTRACT**

The hepatitis C virus (HCV) cell culture system according to the invention consists of human hepatoma cells, which are transfected with a HCV-RNA construct, that comprises 10 the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR as well as a minimum of one marker gene for selection (selection gene).

**Hepatitis C Virus Cell Culture System**DESCRIPTION

5 The invention relates to a hepatitis C virus (HCV) cell culture system, which comprises mainly eukaryotic cells containing transfected HCV specific genetic material, which means they are transfected with HCV specific genetic material.

10

The hepatitis C virus (HCV) is one of the main causes worldwide of chronic and sporadic liver diseases. The history of most HCV infections does not involve any obvious clinical signs, but 80 - 90 % of the infected people become 15 chronic carriers of the virus and 50 % of these chronic carriers of the virus develop chronic hepatitis with different degrees of severity. Approx. 20 % of the chronically infected develop a cirrhosis of the liver over 10 to 20 years, based on what a primary hepatocellular 20 carcinoma can develop. Nowadays chronic hepatitis C is the main indication for liver transplantation. A specific therapy does not exist until now. The only therapy currently available is high-dose administration of Interferon alpha or a combination of Interferon alpha and 25 the purine nucleoside analogue Ribavirin. However, only approx. 60 % of all treated persons respond to this therapy and with these, a new viraemia occurs in more than half of all cases after the discontinuation of the treatment.

30 Due to the high prevalence, especially in industrialized countries, the serious effects of chronic infections and the non-existence of a specific therapy, the development of a HCV specific chemotherapy is an important goal of pharmaceutical research and development. The main problem 35 lies in the previous lack of a suitable cell culture system, which enables the study of virus replication and pathogenesis in eukaryotic cells.

Due to the small amount of virus in blood or tissue, the lack of suitable cell culture systems or animal models (the chimpanzee is still the only possible experimental animal) as well as the lack of efficient systems for producing virus-like particles, it was not possible up to now, to analyze the molecular composition of the HCV particle in-depth and to solve it. The information currently available can be summarized as follows: HCV is an enveloped plus-strand RNA virus with a particle diameter of 50 - 60 nm and a medium density of 1.03 - 1.1 g/ml. It was molecularly cloned and characterized for the first time in 1989 (Choo et al., 1989, *Science*, 244, 359 - 362). The HCV-RNA has a length of approx. 9.6 kb (= 9600 nucleotides), a positive polarity and comprises one open reading frame (ORF), which encodes a linear polyprotein of approx. 3010 amino acids (see Rice 1996, in *Virology*, B. N. Fields, D. M. Knipe, P. M. Howley, Eds. (Lippincott-Raven, Philadelphia, PA, 1996), vol. 1, pp. 931 - 960; Clarke 1997, *J. Gen. Virol.* 78, 2397; and Bartenschlager 1997, *Intervirology* 40, 378 and see Fig. 1 A). During the replication of the virus the polyprotein is cleaved into the mature and functionally active proteins by cellular and viral proteases.

Within the polyprotein the proteins are arranged as follows (from the amino- to the carboxy terminus): Core-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B. The core protein is the main component of the nucleocapsid. The glycoproteins E1 and E2 are transmembrane proteins and the main components of the viral envelope. They probably play an important role during the attachment of the virus to the host cell. These three proteins core, E1, and E2 constitute the viral particle and are therefore called structural proteins. The function of the protein p7 is still not clear. The protein NS2 is probably the catalytic domain of the NS2-3 protease, which is responsible for the processing between the proteins NS2 and NS3. The protein NS3 has two functions, one is a

protease activity in the amino terminal domain, which is essential for the polyprotein processing, and the other a NTPase/helicase function in the carboxy terminal domain, which is probably important during the replication of the 5 viral RNA. The protein NS4A is a co-factor of the NS3 protease. The function of the protein NS4B is unknown.

The open reading frame is flanked on its 5' end by a non-translated region (NTR) approx. 340 nucleotides in length, 10 which functions as the internal ribosome entry site (IRES), and on its 3' end by a NTR approx. 230 nucleotides in length, which is most likely important for the genome replication. A 3' NTR such as this is the object of patent application PCT/US 96/14033. The structural proteins in the 15 amino terminal quarter of the polyprotein are cleaved by host cell signal peptidase. The non-structural proteins (NS) 2 to (NS) 5B are processed by two viral enzymes, namely the NS2-3 and the NS3/4A protease. The NS3/4A protease is required for all cleavages beyond the carboxy 20 terminus of NS3. The function of NS4B is unknown. NS5A, a highly phosphorylated protein, seems to be responsible for the Interferon resistance of various HCV genotypes (see Enomoto et al. 1995, *J. Clin. Invest.* 96, 224; Enomoto et al. 1996, *N. Engl. J. Med.* 334, 77; Gale Jr. et al. 1997, 25 *Virology* 230, 217; Kaneko et al. 1994, *Biochem. Biophys. Res. Commun.* 205, 320; Reed et al., 1997, *J. Virol.* 71, 7187), and NS5B has been identified as the RNA-dependent RNA polymerase.

30 First diagnostic systems have been developed from these findings, which are either based on the detection of HCV specific antibodies in patient serum or the detection of HCV specific RNA using the reverse transcription polymerase chain reaction (RT-PCR), and which are (must be) routinely 35 used with all blood and blood products and/or according to the regulations.

Since the first description of the genome in 1989 several partial and complete sequences of the HCV have been cloned and characterized using the PCR method. A comparison of these sequences shows a high variability of the viral genome in particular in the area of the NS5B gene, which eventually resulted in the classification of 6 genotypes, which are again subdivided into the subtypes a, b, and c.

The genomic variance is not evenly distributed over the genome. The 5'NTR and parts of the 3'NTR are highly conserved, while certain encoded sequences vary a lot, in particular the envelope proteins E1 and E2.

The cloned and characterized partial and complete sequences of the HCV genome have also been analyzed with regard to appropriate targets for a prospective antiviral therapy. In the course of this, three viral enzymes have been discovered, which may provide a possible target. These include (1) the NS3/4A protease complex, (2) the NS3 Helicase and (3) the NS5B RNA-dependent RNA polymerase. The NS3/4A protease complex and the NS3 Helicase have already been crystallized and their three-dimensional structure determined (Kim et al., 1996, Cell, 87, 343; Yem et al., 1998, Protein Science, 7, 837; Love et al., 1996, Cell, 87, 311; Kim et al., 1998, Structure, 6, 89; Yao et al., 1997, Nature Structural Biology, 4, 463, Cho et al., 1998, J. Biol. Chem., 273, 15045). It has not been successful until now with the NS5B RNA-dependent RNA polymerase.

Even though important targets for the development of a therapy for chronic HCV infection have been defined with these enzymes and even though a worldwide intensive search for suitable inhibitors is ongoing with the aid of rational drug design as well as high throughput screening, the development of a therapy has one major deficiency, namely the lack of cell culture systems or simple animal models, which allow direct, reliable identification of HCV-RNA or

HCV antigens with simple methods which are common in the laboratory. The lack of these cell culture systems is also the main reason that to date the comprehension of HCV replication is still incomplete and mainly hypothetical.

5

Although according to the experts, a close evolutionary relationship exists between HCV and the flavivirus and pestiviruses, and self-replicating RNAs have been described for these, which can be used for the replication in different cell lines with a relatively high yield, (see Khromykh et al., 1997, *J. Virol.* 71, 1497; Behrens et al., 1998, *J. Virol.* 72, 2364; Moser et al., 1998, *J. Virol.* 72, 5318), similar experiments with HCV have not been successful to date.

15

Although it is known from different publications that cell lines or primary cell cultures can be infected with high titre patient serum containing HCV, (Lanford et al. 1994, *Virology* 202, 606; Shimizu et al. 1993, *Proceedings of the National Academy of Sciences, USA*, 90, 6037 - 6041; Mizutani et al. 1996, *Journal of Virology*, 70, 7219 - 7223; M. Ikeda et al. 1998, *Virus Res.* 56, 157; Fournier et al. 1998, *J. Gen. Virol.* 79, 2376 and bibliographical references quoted in here; Ito et al. 1996, *Journal of General Virology*, 77, 1043 - 1054), these virus-infected cell lines or cell cultures do not allow the direct detection of HCV-RNA or HCV antigens. The viral RNA in these cells can not be detected in a Northern Blot (a standard method for the quantitative detection of RNA) or the viral protein in a Western Blot or with immunoprecipitation. It has only been possible to detect HCV replication with very costly and indirect methods. These disadvantageous facts show that obviously the replication in these known virus-infected cell lines or cell cultures is completely insufficient.

Furthermore it is known from the publications of Yoo et al. (1995, *Journal of Virology*, 69, 32 - 38) and of Dash et al., (1997, *American Journal of Pathology*, 151, 363 - 373) that hepatoma cell lines can be transfected with synthetic HCV-RNA, which are obtained through in vitro transcription of the cloned HCV genome. In both publications the authors started from the basic idea that the viral HCV genome is a plus-strand RNA functioning directly as mRNA after being transfected into the cell, permitting the synthesis of viral proteins in the course of the translation process, and so new HCV particles are (could be) formed. This viral replication, which means these newly formed HCV viruses and their RNA, have been detected through RT-PCR. However the published results of the RT-PCR carried out indicate, that the HCV replication in the described HCV transfected hepatoma cells is not particularly efficient and is not sufficient to measure the quality, let alone the quantity of the fluctuations in the replication rate after an targeted action with prospective antiviral treatments.

Furthermore it is prior art (Yanagi et al., Proc. Natl. Acad. Sci. USA, 96, 2291-95, 1999), that the highly conserved 3' NTR is essential for the virus replication. This knowledge strictly contradicts the statements of Yoo et al. and Dash et al., who used for their experiments only HCV genomes with shorter 3' NTRs since they did not know the authentic 3' end of the HCV genome.

The object of the present invention is to provide a HCV cell culture system, where the viral RNA self-replicates in the transfected cells with such a high efficiency that the quality and quantity of the fluctuations in the replication rate can be measured with common methodologies usually found in the laboratory after a targeted action with virus and prospective HCV specific antivirals in particular.

35

The solution to this problem is to provide a cell culture system of the prior mentioned type, where the eukaryotic

- cells are human cells, in particular hepatoma cells, which are preferably derived from a normal hepatoma cell line, but can also be obtained from an appropriate primary cell culture, and where the transfected HCV specific genetic material is a HCV-RNA construct, which essentially comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR preferably in the order mentioned as well as a minimum of one marker gene for selection (selection gene).
- 10 Here and in the following "NTR" stands for "non-translated region" and is a known and familiar term or abbreviation to the relevant expert.
- 15 Here and in the following the term "HCV-RNA construct" comprises constructs, which include the complete HCV genome, as well as those, which only include a part of it, which means a HCV subgenome.
- 20 A preferred variation of the cell culture system according to the invention, which had proven to be worthwhile in practice, is lodged at the DSMZ, Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (German collection of Microorganisms and Cell Cultures) in Braunschweig, Germany  
25 under the number DSM ACC2394 (laboratory name HuBl 9-13).
- With the cell culture system according to the invention an in vitro system is provided for the first time, where HCV-RNA is self-replicated and expressed intracellularly and in a sufficient amount, so that the quantity of the amounts of HCV- RNA as well as the HCV specific proteins can be determined with conventional and reliably precise biochemical measuring methods. This means an almost authentic cell-based HCV replication system is available  
30 for the first time, which is urgently needed for the development and testing of antiviral drugs. This test system provides the possibility of identifying potential  
35

targets for an effective HCV specific therapy and developing and evaluating HCV specific chemotherapeutics.

The invention is based on the surprising finding that efficient replication of the HCV-RNA only occurs in cells if they have been transfected with an HCV-RNA construct, which comprises at least the 5' and the 3' non-translated regions (NTR) and the non-structural proteins (NS) 3 to 5B and additionally a marker gene for selection (selection gene). The structural genes are obviously without great importance for replication, whereas efficient replication of the HCV-RNA apparently only occurs if the transfected cells are subject to permanent selection pressure, which is imparted by the marker gene for selection (selection gene) linked to the HCV-RNA. Consequently the marker gene (selection gene) seems on one hand to provoke the selection of those cells, where the HCV-RNA replicates productively, and it seems on the other hand to considerably increase the efficiency of the RNA replication.

An object of the invention is also a cell-free HCV-RNA construct, characterized in that it comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR, preferably in the order mentioned, as well as a marker gene for selection (selection gene).

In the present context the terms 5' NTR and NS3 and NS4A and NS4B and NS5A and NS5B and 3' NTR comprise each nucleotide sequence, which is described in the state of the art as the nucleotide sequence for each functional segment of the HCV genome.

By providing a HCV-RNA construct such as this, a detailed analysis of the HCV replication, pathogenesis and evolution in cell culture is possible for the first time. The HCV specific viral RNA can specifically be created as a complete genome or subgenome in any amount, and it is

possible, to manipulate the RNA construct and consequently to examine and identify the HCV functions on a genetic level.

5 Because all HCV enzymes identified as a main target for a therapy at the moment, namely the NS3/4A protease, the NS3 helicase and the NS5B polymerase, are included in the HCV-RNA construct according to the invention, it can be used for all relevant analyses.

10 An embodiment of the HCV-RNA construct, which has proven to be worthwhile in practical use, stands out by the fact that it comprises the nucleotide sequence according to the sequence protocol SEQ ID NO:1.

15 Further embodiments with similar good properties for practical use are characterized in that they comprise a nucleotide sequence either according to sequence protocol SEQ ID NO:2 or according to sequence protocol SEQ ID NO:3  
20 or according to sequence protocol SEQ ID NO:4 or according to sequence protocol SEQ ID NO:5 or according to sequence protocol SEQ ID NO:6 or according to sequence protocol SEQ ID NO:7 or according to sequence protocol SEQ ID NO:8 or according to sequence protocol SEQ ID NO:9 or according to  
25 sequence protocol SEQ ID NO:10 or according to sequence protocol SEQ ID NO:11.

It is possible to provide the HCV subgenomic construct with a 3' NTR, which has a nucleotide sequence so far unknown in  
30 the state of art, a nucleotide sequence, which has been selected from the group of nucleotide sequences (a) to (i) listed in the following:

(a) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTAGCTT  
35 TTTTTTTCTTTTTTGAGAGAGAGTCTCACCTGTGCCAGACTGGAG  
T

(b) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT TTTTTAGTCT  
TTTTTTTTTC TTTTTTTGA GAGAGAGAGT CTCACTCTGT TGCCCAGACT  
GGAGC

5 (c) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT TTTAATCTTT  
TTTTTTTCT TTTTTTTGA GAGAGAGAGT CTCACTCTGT TGCCCAGACT  
GCAGC

10 (d) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT TTTTTAGTC  
TTTTTTTTT TCTTTTTTT TGAGAGAGAG AGTCTCACTC TGTTGCCAG  
ACTGGAGT

15 (e) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT TTTTTAGTCT  
TTTTTTTTT TCTTTTTTT TGAGAGAGAG AGTCTCACTC TGTTGCCAG  
ACTGGAGT

(f) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT TTTTTAGTCT  
TTTTTTTTT TCTTTTTTT TTGAGAGAGA GAGTCTCACT CTGTTGCCCA  
GACTGGAGT

20 (g) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT TTTTTAGTCT  
TTTTTTTTT CTTTTTTTT GAGAGAGAGA GTCTCACTCT GTTGCCAGA  
CTGGAGT

25 (h) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT TTTTTTAAT  
CTTTTTTTT TTTTCCTTT TTTTGAGAGA GAGAGTCTCA CTCTGTTGCC  
CAGACTGGAG T

(i) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTT TTTTTTAATC  
30 TTTTTTTTT TTTTCTTTT TTTTGAGAG AGAGAGTCTC ACTCTGTTGC  
CCAGACTGGA GT

35 The marker gene for selection (selection gene) included in  
the HCV-RNA constructs according to the invention is  
preferably a resistance gene, in particular an antibiotic  
resistance gene.

This has the advantage that the cells transfected with this construct can easily be selected from the non-transfected cells by adding for example the appropriate antibiotic to the cell culture medium in the case of an antibiotic resistance gene.

In the present context 'antibiotic' means any substance, which impedes the non-transfected host cells or the cells, where the HCV-RNA is not replicating efficiently, continuing to live or grow, especially the cell poison Puromycin, Hygromycin, Zeocin, Bleomycin or Blasticidin.

A preferred marker gene for selection (selection gene) and resistance gene, which has proven to be worthwhile in practice, is the neomycin phosphotransferase gene.

An alternative for the antibiotic resistance genes is for example the thymidine kinase gene, which can be used to carry out a HAT selection.

The marker gene for selection (selection gene), the preferred resistance gene and the most preferred antibiotic resistance gene is preferably positioned in the HCV-RNA construct after the HCV 5' NTR, which means downstream from the 5' NTR and upstream from the HCV reading frame. However, an insertion in the area of the 3' NTR or another site of the HCV genome or subgenome, for example within the polyprotein, is also conceivable.

In another embodiment of the HCV-RNA construct according to the invention the marker gene for selection (selection gene), in particular an antibiotic resistance gene, is linked to the HCV-RNA or HCV genomic or subgenomic sequence via a ribozyme or a recognition site for a ribozyme.

This has the advantage, that after the selection of the cells, in which the HCV-RNA is replicating productively,

the resistance gene in the obtained cell clones can be separated from the HCV subgenomic sequence through a ribozyme-dependent cleavage, namely by activating the inserted ribozyme or in the case of a construct with a 5 recognition site for a ribozyme, by transfecting the ribozyme into the cells (for example through the transfection of a ribozyme construct or infection with a viral expression vector, into which the appropriate ribozyme has been inserted). By this means an authentic HCV 10 genomic construct can be obtained without a resistance gene, which can then form authentic infectious virus particles.

Another preferred embodiment of the HCV-RNA construct 15 according to the invention is characterized in that the construct has at least one integrated reporter gene.

In the following a reporter gene means any gene, whose presence can be easily detected with, in general, simple 20 biochemical or also histochemical methods after being transferred into a target organism, which means a gene, that encodes for a protein, which can be easily and reliably detected and quantified in small amounts with the common measuring methods in the laboratory.

25 This variation of the HCV-RNA construct has the advantage that the extent of the replication of this construct can be easily and quickly measured with the methods common in the laboratory using the reporter gene product.

30 The reporter gene is preferably a gene from the group of the luciferase genes, the CAT gene (chloramphenicol acetyl transferase gene), the lacZ gene (beta galactosidase gene), the GFP gene (green fluorescence protein gene), the GUS 35 gene (glucuronidase gene) or the SEAP gene (secreted alkaline phosphatase gene). This reporter gene and its products, namely the relevant reporter proteins, can be

detected for example using fluorescence, chemiluminescence, colorimetric measurements or by means of immunological methods (for example ELISA).

5 A surrogate marker gene can also be considered as a reporter gene. In this context it includes those genes, which encode for cellular proteins, nucleic acids or generally for those functions, which are subject to variation depending on the replication of the virus, and  
10 which consequently are either suppressed or activated in the cells, in which the HCV or the HCV-RNA construct multiplies. This means, the suppression or activation of this function is a surrogate marker for the replication of the virus or the replication of the HCV-RNA construct.

15 The positions of the reporter genes and the marker gene for selection (selection gene) can be selected in such a way, that a fusion protein made from both genetic products will be expressed. This has the advantage that these two genes  
20 can be arranged in such a way in the HCV-RNA construct that their two expressed proteins are fused via a recognition sequence for a protease (for example ubiquitin) or via a self-cleaving peptide (for example the 2A protein of the Picornaviruses) at first and will be separated  
25 proteolytically later.

These two positions might as well lie apart from each other in such way, that both genetic products are separately expressed (for example in the order: marker or resistance  
30 gene – internal ribosome binding site – reporter gene).

In the case of the reporter gene one embodiment has proven to be particularly worthwhile, where the reporter gene is cloned into the open reading frame of the HCV genome or  
35 subgenome in such a way that it will only be transferred to an active form after proteolytic processing.

The cell culture system according to the invention can be used for various purposes in each of its embodiments. These comprise:

- 5     • The detection of antiviral substances. This can include for example: organic compounds, which interfere directly or indirectly with viral growth (for example inhibitors of the viral proteases, the NS3 helicase, the NS5B RNA-dependent RNA polymerase), antisense oligonucleotides, 10 which will hybridize to any target sequence in the HCV-RNA construct (for example the 5' NTR) and will have an direct or indirect influence on the virus growth for example due to a reduction of translation of the HCV polyprotein or ribozymes, which cleave any HCV-RNA 15 sequence and consequently impair virus replication.
- The evaluation of any type of antiviral substances in the cell culture. These substances can be detected on the isolated purified enzyme for example with 'rational drug design' or 'high-throughput screening'. Evaluation means mainly the determination of the inhibitory features of the respective substance as well as its mode of action. 20
- 25     • The identification of new targets of viral or cellular origin for a HCV specific antiviral therapy. If for example a cellular protein is essential for viral replication, the viral replication can also be influenced by inhibiting this cellular protein. The system according to the invention also enables the detection of these auxiliary factors. 30
- 35     • The determination of drug resistance. It can be assumed that resistance to therapy occurs due to the high mutation rate of the HCV genome. This resistance, which is very important for the clinical approval of a substance, can be detected with the cell culture system

according to the invention. Cell lines, in which the HCV-RNA construct or the HCV genome or subgenome replicates, are incubated with increasing concentrations of the relevant substance and the replication of the viral RNA is either determined by means of an introduced reporter gene or through the qualitative or quantitative detection of the viral nucleic acids or proteins. Resistance is given if no or a reduced inhibition of the replication can be observed with the normal concentration of the active substance. The nucleotide and amino acid replacements responsible for the therapy resistance can be determined by recloning the HCV-RNA (for example by the means of RT-PCR) and sequence analysis. By cloning the relevant replacement/s into the original construct its causality for the resistance to therapy can be proven.

- The production of authentic virus proteins (antigens) for the development and/or evaluation of diagnostics.  
The cell culture system according to the invention also allows the expression of HCV antigens in cell cultures. In principle these antigens can be used as the basis for diagnostic detection methods.
- The production of HCV viruses and virus-like particles, in particular for the development or production of therapeutics and vaccines as well as for diagnostic purposes. Especially cell culture adapted complete HCV genomes, which could be produced by using the cell culture system according to the invention, are able to replicate in cell culture with high efficiency. These genomes have the complete functions of HCV and in consequence they are able to produce infectious viruses.
- The HCV-RNA construct according to the invention by itself can also be used for various purposes in all its embodiments. This includes first of all:

- The construction of attenuated hepatitis C viruses or HCV-like particles and their production in cell cultures:

5

Attenuated HCV or HCV-like particles can be created by accidental or purposefully introduced mutations, such as point mutations, deletions or insertions, which means viruses or virus-like particles with complete ability to replicate, but reduced or missing pathogenicity. These attenuated HCV or HCV-like particles can be used in particular as vaccine.

10

- The construction of HCV-RNA constructs with integrated foreign genes, used for example as liver cell specific vector in gene therapy. Due to the distinctive liver cell tropism of the HCV and the possibility of replacing parts of the genome by heterologous sequences, HCV-RNA constructs can be produced, where for example the structural proteins can be replaced by a therapeutically effective gene. The HCV-RNA construct obtained in this way is introduced into cells preferably by means of transfection, which express the missing HCV functions, for example the structural proteins, in a constitutive or inducible way. Virus particles, carrying the HCV-RNA construct, can be created by means of this method known to the expert under the term 'transcomplementation'. The particles obtained can preferably be used for the infection of liver cells. Within these the therapeutically effective foreign gene will be expressed and will consequently develop its therapeutic effect.

25

30

- The detection of permissive cells, which means cells, in which a productive virus growth occurs. For this purpose either one of the HCV-RNA genomic constructs previously mentioned, which is able to form complete infectious viruses, or one of the HCV subgenomes previously

mentioned, which according to the previously mentioned example will be transfected in a cell line first, which expresses the missing functions in a constitutive or inducible way, is used. In each case virus particles are  
5 created, which carry a resistance and/or reporter gene apart from the HCV sequence. In order to detect cells, where the HCV is able to replicate, these cells are infected with viruses generated in this way and subject to an antibiotic selection or they are examined  
10 depending on the HCV-RNA construct by means of determining the presence of the expression of the reporter gene. Because an antibiotic resistance or reporter gene expression can only be established, when HCV-RNA construct replicates, the cells detected in this  
15 way must be permissive. Almost any cell line or primary cell culture can be tested in regard to the permissivity and detected in this way.

- 20 The cell culture system according to the invention also permits targeted discovery of HCV-RNA constructs for which there is an increase in the efficiency of replication due to mutations. This occurs either by chance, in the context of HCV-RNA replication, or by targeted introduction into  
25 the construct. These mutations, leading to a change in the replication of the HCV-RNA construct, are known to experts as adaptive mutations." The invention therefore also includes a method for obtaining cell culture adapted mutants of a HCV-RNA construct according to the invention  
30 following the above description, in which the mutants have increased replication efficiency compared to the original HCV-RNA construct. It further includes a method for the production of mutants of a HCV-RNA full-length genome or of a HCV-RNA subgenome or of any HCV-RNA construct with  
35 increased replication efficiency compared to the original HCV-RNA full-length genome or subgenome or HCV-RNA construct, as well as cell culture adapted mutants of HCV-

RNA constructs, HCV-RNA full-length genomes and HCV subgenomes with increased replication efficiency compared to the original constructs, subgenomes or full-length genomes.

5

The method according to the invention for the production of cell culture adapted mutants of a HCV-RNA construct according to the invention, in which the mutants have increased replication efficiency compared to the HCV-RNA construct, is characterised in that a cell culture system according to claim 1, in which the transfected HCV specific genetic material is a HCV-RNA construct with a selection gene according to claims 4 to 19, is cultivated on/in the selection medium corresponding to the selection gene, that 15 the cultivated cell clones are collected and that the HCV-RNA construct is isolated from these cell clones.

In an advantageous extension of this production method, the 20 isolated HCV-RNA constructs are passaged at least one more time, that is they are transfected in cells of a cell culture system according to claim 1, the thus obtained cell culture system according to claim 1, in which the transfected HCV specific genetic material is the isolated HCV-RNA construct with a selection gene, is cultivated 25 on/in the selection medium corresponding to the selection gene, the cultivated cell clones are collected and the HCV-RNA constructs are thus isolated.

Using this process variation, the quantity of adaptive 30 mutations and hence the degree of replication efficiency in the relevant HCV-RNA constructs can be increased even further.

The method according to the invention for the production of 35 mutants of a HCV-RNA full-length genome or of a HCV-RNA subgenome or of any HCV-RNA construct with increased replication efficiency compared to the original HCV-RNA

full-length genome or subgenome or HCV-RNA construct, has the following features. Using one of the two production methods presented above, a cell culture adapted mutant of a HCV-RNA construct is produced, isolated from the cells,  
5 cloned using state of the art known methods and sequenced. By comparing with the nucleotide and amino acid sequence of the original HCV-RNA construct, the type, number and position of the mutations is determined. These mutations are then introduced into an (isolated) HCV subgenome or  
10 full-length genome or any HCV-RNA construct, either by site-directed mutagenesis, or by exchange of DNA fragments containing the relevant mutations.

A test can be carried out to determine or verify which  
15 mutations actually are responsible for an alteration of replication efficiency, particularly an increase in replication. In this test the corresponding nucleotide and/or amino acid changes are introduced into the original HCV-RNA construct and the modified construct is then  
20 transfected in cell culture. If the introduced mutation actually leads to an increase in replication, then for a HCV-RNA construct with a selectable marker gene, the number of resistant cell clones in the artificially mutated construct should be noticeably higher compared to the  
25 untreated construct.

In the case of a construct with a reporter gene, the activity or quantity of the reporter should be noticeably higher for the artificially mutated construct compared to  
30 the untreated one.

The cell culture adapted HCV-RNA constructs with high replication efficiency according to the invention are characterized in that, through nucleotide or amino acid exchanges, they are derivable from a HCV-RNA construct according to one of the claims 4 to 19, and that they are  
35

obtainable using one of the two production processes presented above.

These cell culture adapted HCV-RNA constructs can be used  
5 to produce any HCV-RNA constructs or HCV full-length or  
subgenomes with increased replication efficiency. Both  
constructs with a selectable resistance gene and constructs  
without one or with a non-selectable reporter gene (e.g.  
10 luciferase) can be produced in this way, since replication  
of cell culture adapted HCV-RNA constructs can also be  
demonstrated in non-selected cells due to their high  
replication efficiency.

The cell culture adapted mutants of a HCV-RNA construct or  
15 HCV-RNA full-length genome or HCV subgenome with high  
replication efficiency compared to the original HCV-RNA  
construct or the original HCV full-length genome are  
characterized in that they are obtainable by a method in  
which the type and number of mutations in a cell culture  
20 adapted HCV-RNA construct are determined through sequence  
analysis and sequence comparison and these mutations are  
introduced into a HCV-RNA construct, particularly a HCV-RNA  
construct according to one of the claims 4 to 19, or into  
an (isolated) HCV-RNA full-length genome, either by site-  
25 directed mutagenesis, or by exchange of DNA fragments  
containing the relevant mutations.

A group of preferred HCV-RNA constructs, HCV full-length  
genomes and HCV subgenomes with high and very high  
30 replication efficiency, which are consequently highly  
suitable for practical use is characterised in that it  
contains one, several or all of the amino acid or nucleic  
acid exchanges listed in table 3 and/or one or several of  
the following amino acid exchanges: 1283 arg -> gly , 1383  
35 glu -> ala , 1577 lys -> arg , 1609 lys -> glu , 1936  
pro -> ser , 2163 glu -> gly , 2330 lys -> glu , 2442 ile

-> val. (The numbers refer to the amino acid positions of the polyprotein of the HCV isolate con1, see Table 1).

5 Special features of the nucleotid sequences according to the sequence listings:

SEQ ID-NO: 1

Name: I389/Core-3'/wt

Composition (Nucleotide positions):

1.1-341: HCV 5' non-translated region

10 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker

3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame

15 4.1813-10842: HCV Polyprotein from Core up to nonstructural protein 5B

5. 1813-2385: HCV Core Protein; structural protein

6. 2386-2961: envelope protein 1 (E1); structural protein

20 7. 2962-4050: envelope protein 2 (E2); structural protein

8. 4051-4239: Protein p7

9. 4240-4890: nonstructural protein 2 (NS2); HCV NS2-3 Protease

25 10.4891-6783: nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase

11.6784-6945: nonstructural protein 4A (NS4A); NS3 Protease cofactor

12.6946-7728: nonstructural protein 4B (NS4B)

30 13.7729-9069: nonstructural protein 5A (NS5A)

14.9070-10842: nonstructural protein 5B (NS5B); RNA-dependent RNA-polymerase

15.10846-11076: HCV 3' non-translated region

- 22 -

**SEQ ID-NO: 2**

Name: I337/NS2-3'/wt

Composition (Nucleotide positions):

1.1-341: HCV 5' non-translated region

5 2.342-1181: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker

3.1190-1800: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame

10 4.1801-8403: HCV Polyprotein from nonstructural protein 2 up to nonstructural protein 5B

5. 1801-2451: nonstructural protein 2 (NS2); HCV NS2-3 Protease

6. 2452-4344: nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase

7. 4345-4506: nonstructural protein 4A (NS4A); NS3 Protease cofactor

8. 4507-5289: nonstructural protein 4B (NS4B)

9. 5290-6630: nonstructural protein 5A (NS5A)

20 10. 6631-8403: nonstructural protein 5B (NS5B); RNA-dependent RNA-polymerase

11.8407-8637: HCV 3' non-translated region

**SEQ ID-NO: 3**

25 Name: I389/NS3-3'/wt

Composition (Nucleotide positions):

1.1-341: HCV 5' non-translated region

2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker

30 3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame

4.1813-7767: HCV Polyprotein from nonstructural protein 3 up to nonstructural protein 5B

5. 1813-3708: nonstructural protein 3 (NS3); HCV NS3  
Protease/Helicase
6. 3709-3870: Nonstructural protein 4A (NS4A); NS3  
Protease Cofactor
- 5 7. 3871-4653: Nonstructural protein 4B (NS4B)
8. 4654-5994: Nonstructural protein 5A (NS5A)
9. 5995-7767: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase
10. 7771-8001: HCV 3' non-translated Region

10

**SEQ ID-NO: 4**

Name: I337/NS3-3'/wt

Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region
- 15 2. 342-1181: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
3. 1190-1800: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
- 20 4. 1801-7758: HCV Polyprotein from Nonstructural protein 3 up to Nonstructural protein 5B
5. 1801-3696: Nonstructural protein 3 (NS3); HCV NS3  
Protease/Helicase
6. 3697-3858: Nonstructural protein 4A (NS4A); NS3  
Protease Cofactor
- 25 7. 3859-4641: Nonstructural protein 4B (NS4B)
8. 4642-5982: Nonstructural protein 5A (NS5A)
9. 5983-7755: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase
- 30 10. 7759-7989: HCV 3' non-translated Region

**SEQ ID-NO: 5**

Name: I389/NS2-3'/wt

Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region
- 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
- 3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
- 4.1813-8418: HCV Polyprotein from Nonstructural protein 2 up to Nonstructural protein 5B
5. 1813-2463: Nonstructural protein 2 (NS2); HCV NS2-3 Protease
6. 2464-4356: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
7. 4357-4518: Nonstructural protein 4A (NS4A); NS3 Protease Cofactor
- 15 8. 4519-5301: Nonstructural protein 4B (NS4B)
9. 5302-6642: Nonstructural protein 5A (NS5A)
10. 6643-8415: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase
11. 8419-8649: HCV 3' non-translated Region

20

**SEQ ID-NO: 6**

Name: I389/NS3-3'/9-13F

Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region
- 25 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
- 3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
- 30 4.1813-7767: HCV Polyprotein from Nonstructural protein 3 up to Nonstructural protein 5B of the cell culture-adapted mutant 9-13F
5. 1813-3708: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
- 35 6. 3709-3870: Nonstructural protein 4A (NS4A); NS3

- 25 -

## Protease Cofactor . . .

7. 3871-4653: Nonstructural protein 4B (NS4B)
  8. 4654-5994: Nonstructural protein 5A (NS5A)
  9. 5995-7767: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase
- 5 7771-8001: HCV 3' non-translated Region

**SEQ ID-NO: 7**

Name: I389/Core-3'/9-13F

## 10 Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region
- 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
- 3.1202-1812: internal ribosome entry site from
- 15 encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
- 4.1813-10842: HCV Polyprotein from Core up to Nonstructural protein 5B of the cell culture-adapted mutant 9-13F
5. 1813-2385: HCV Core Protein; structural protein
- 20 6. 2386-2961: envelope protein 1 (E1); structural protein
7. 2962-4050: envelope protein 2 (E2); structural protein
8. 4051-4239: Protein p7
- 25 9. 4240-4890: Nonstructural protein 2 (NS2); HCV NS2-3 Protease
- 10.4891-6783: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
- 11.6784-6945: Nonstructural protein 4A (NS4A); NS3
- 30 Protease Cofactor
- 12.6946-7728: Nonstructural protein 4B (NS4B)
- 13.7729-9069: Nonstructural protein 5A (NS5A)
- 14.9070-10842: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase
- 35 15.10846-11076: HCV 3' non-translated Region

**SEQ ID-NO: 8**

Name: I389/NS3-3'/5.1

Composition (Nucleotide positions):

1.1-341: HCV 5' non-translated region

5 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker

3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame

10 4.1813-7767: HCV Polyprotein from Nonstructural protein 3 up to Nonstructural protein 5B of the cell culture-adapted mutant 5.1

5. 1813-3708: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase

15 6. 3709-3870: Nonstructural protein 4A (NS4A); NS3 Protease Cofactor

7. 3871-4653: Nonstructural protein 4B (NS4B)

8. 4654-5994: Nonstructural protein 5A (NS5A)

9. 5995-7767: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase

20 7771-8001: HCV 3' non-translated Region

**SEQ ID-NO: 9**

Name: I389/Core-3'/5.1

25 Composition (Nucleotide positions):

1.1-341: HCV 5' non-translated region

2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker

3.1202-1812: internal ribosome entry site from

30 encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame

4.1813-10842: HCV Polyprotein from Core up to Nonstructural protein 5B of the cell culture-adapted mutant 5.1

5. 1813-2385: HCV Core Protein; structural protein

6. 2386-2961: envelope protein 1 (E1); structural protein
7. 2962-4050: envelope protein 2 (E2); structural protein
- 5 8. 4051-4239: Protein p7
9. 4240-4890: Nonstructural protein 2 (NS2); HCV NS2-3 Protease
- 10 10. 4891-6783: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
11. 6784-6945: Nonstructural protein 4A (NS4A); NS3 Protease Cofactor
12. 6946-7728: Nonstructural protein 4B (NS4B)
13. 7729-9069: Nonstructural protein 5A (NS5A)
14. 9070-10842: Nonstructural protein 5B (NS5B); RNA-  
15 dependent RNA-Polymerase
15. 10846-11076: HCV 3' non-translated Region

**SEQ ID-NO: 10**

Name: I389/NS3-3'/19

- 20 20. Composition (Nucleotide positions):
  - 1.1-341: HCV 5' non-translated region
  - 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
  - 3.1202-1812: internal ribosome entry site from  
25 encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
  - 4.1813-7767: HCV Polyprotein from Nonstructural protein 3 up to Nonstructural protein 5B of the cell culture-adapted mutant 19
- 30 5. 1813-3708: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
6. 3709-3870: Nonstructural protein 4A (NS4A); NS3 Protease Cofactor
7. 3871-4653: Nonstructural protein 4B (NS4B)
- 35 8. 4654-5994: Nonstructural protein 5A (NS5A)

9. 5995-7767: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase  
7771-8001: HCV 3' non-translated Region

## 5 SEQ ID-NO: 11

Name: I389/Core-3'/19

Composition (Nucleotide positions):

1.1-341: HCV 5' non-translated region

2.342-1193: HCV Core Protein-Neomycin Phosphotransferase

10 fusion protein; selectable Marker

3.1202-1812: internal ribosome entry site from  
encephalomyokarditis virus; directs translation of the  
downstream located HCV open reading frame

4.1813-10842: HCV Polyprotein from Core up to Nonstructural  
15 protein 5B of the cell culture-adapted mutant 19

5. 1813-2385: HCV Core Protein; structural protein

6. 2386-2961: envelope protein 1 (E1); structural  
protein

7. 2962-4050: envelope protein 2 (E2); structural  
20 protein

8. 4051-4239: Protein p7

9. 4240-4890: Nonstructural protein 2 (NS2); HCV NS2-3  
Protease

10.4891-6783: Nonstructural protein 3 (NS3); HCV NS3  
25 Protease/Helicase

11.6784-6945: Nonstructural protein 4A (NS4A); NS3  
Protease Cofactor

12.6946-7728: Nonstructural protein 4B (NS4B)

13.7729-9069: Nonstructural protein 5A (NS5A)

30 14.9070-10842: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase

15.10846-11076: HCV 3' non-translated Region

The invention is described in detail in the following by way of examples of embodiments and respective diagrams. These are as follows

5 **Fig. 1 A:** The structure of a HCV-RNA construct according to the invention

10 On the top a diagram of the structure of the complete parental HCV genome is given with the positions of the genes for the cleavage products core, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B within the polyprotein, and the 5' and 3' non-translated regions (5' NTR and 3' NTR) - shown as a thick horizontal line -, and with the 15 two positions selected for the creation of the subgenomic constructs, namely the position of the 'GDD catalytic domain' of the NS5B RNA polymerase (GDD) and the position of the 3' end of the HCV-IRES (nucleotide positions 1 to 377 and 1 to 389) - drawn above the diagram of the genome. The 20 numbers below the diagram of the genome indicate the respective nucleotide positions.

25 Diagrams of the structure of two modified HCV-RNA constructs (subgenome) according to the invention are shown below, consisting of the 5' HCV-IRES, the neomycin phosphotransferase gene (Neo<sup>R</sup>), the EMCV-IRES (E-I) and the HCV sequences of NS2 or 30 NS3 up to the authentic 3' end. The position of the 10-amino acid deletion comprising the NS5B polymerase GDD motive is marked with a triangle (Δ).

35 **Fig. 1 B:** The result of a denaturing formaldehyde-agarose gel electrophoresis for the detection of replicating plus-strand RNA in transfected subpassaged Huh 7 cell clones.

5 The positions of HCV specific RNAs (arrows) and the 28S rRNA are specified to the right of lane 12, the size (number of nucleotides) of the RNA marker (M) is specified to the left of lane 1.

10 **Fig. 1 C:** The result of a PCR test with subsequent Southern Blot to demonstrate the absence of integrated replicon DNA in most of the selected cell clones.

15 The lanes 1 and 2 show the positive control, lane 13 the negative control. The figures to the left of lane 1 indicate the size of the nucleotide marker molecules.

15 **Fig. 2 A:** The result of a PCR test with subsequent Southern Blot for the detection of the sensitive exclusion of integrated replicon DNA (plasmid molecule I<sub>377</sub>/NS3-3'/wt) in a cell clone containing a HCV-RNA construct (9-13).

25 The lanes 7 to 11 represent the result of a titration of DNA molecules of the construct I<sub>377</sub>/NS3-3'/wt without addition of total DNA of the cell clone 9-13, and the lanes 2 - 6 represent the result obtained with the same plasmid molecules with the addition of 1 µg 9-13 DNA each prior to the PCR (for the purpose of excluding an inhibitor of the PCR in the DNA preparation). Lane 13 represents the negative control (PCR without DNA template). Lane 1 shows the result, which was achieved with one µg total DNA of the cell clone 9-13.

30 **Fig. 2 B:** The result of a Northern Blot test to quantify the amounts of HCV plus- and minus-strand RNA.

The arrows mark the positions of replicon RNA.

5 The "plus" and "minus" details indicate the positive (plus) and negative (minus) polarity of the RNA controls, which have been applied to the gel. "Minus-strand" and "Plus-strand" indicate the specificity of the radioactive RNA probes.

10 **Fig. 2 C:** The result of a formaldehyde-agarose gel electrophoresis after radioactive labeling of the intracellular replicated HCV-RNA to demonstrate the resistance of HCV-RNA replication to dactinomycin.

15 **Fig. 3 A:** The detection of HCV specific antigens in the selected cell clones by means of immunoprecipitation after metabolic radioactive labeling.

20 The lanes 7 - 9 represent authentic size marker (which have been obtained in Huh 7 cells after the transient expression of a HCV-RNA construct); identified HCV proteins are labeled on the left edge of lane 1, the molecular weights (in Kilodalton) are specified to the right of lane 9.

25 **Fig. 3 B:** Results of an immunofluorescence test to establish the subcellular localisation of HCV antigens.

30 **Fig. 4 :** Diagram of the structure of a selectable HCV-RNA construct according to the invention (complete genome) consisting of the 5' HCV-IRES, the neomycin phosphotransferase gene (NeoR), a heterologous IRES element, for example from the encephalomyocarditis virus (E-I), the complete HCV open reading frame and the authentic 3' NTR.

5           **Fig. 5:** Diagram of the structure of HCV-RNA constructs  
with inserted antibiotic resistance gene **(A)**  
within the nucleotide sequence encoding the  
polyprotein (monocistronic RNA construct), and  
**(B)** within the 3' NTR (bicistronic RNA  
construct).

10           **Fig. 6:** Diagram of the structure of HCV-RNA constructs  
with inserted reporter genes **(A)** as part of a HCV  
replicon from NS3 to NS5B, - in the end the  
reporter protein is cleaved by viral or cellular  
proteases out of the polyprotein or the marker  
15           gene for selection (selection gene) or the  
resistance gene are transfected into the cells  
through co-transfection -, **(B)** as part of a  
fusion gene composed of a resistance and reporter  
gene (for example for the neomycin  
20           phosphotransferase and green fluorescent  
protein), **(C)** as part of a replicon composed of a  
resistance and reporter gene (for example the  
neomycin phosphotransferase and green fluorescent  
protein) bound via a nucleotide sequence, which  
encodes for an amino acid sequence (area with  
25           hatches), which can be cleaved by a protease or  
has a self-cleaving (autocatalytic) activity, **(D)**  
as independent gene (in this case the green  
fluorescent protein), which is expressed from its  
own internal ribosome binding site (IRES); - the  
30           resistance gene (in this case the neomycin  
phosphotransferase gene) is also expressed from  
its own internal ribosome binding site (IRES)  
(polycistronic construct).

35           **Fig. 7:** Diagram of the structure of a HCV-RNA construct,  
where the resistance gene is linked to the HCV-

RNA sequence via a ribozyme or a recognition site for a ribozyme.

The thick lines illustrate the HCV 5' and 3' NTRs, E-I is a heterologous internal ribosome binding site, which is required for the expression of the resistance gene, and the grey square illustrates the ribozyme or a recognition site for a ribozyme.

**Fig. 8:** Diagram of the structure of a HCV-RNA construct with resistance gene and integrated foreign gene.

**Fig. 9:** Method for comparing the specific infectivity (expressed as number of cell colonies obtained) of total RNA against in vitro transcripts. HCV-RNA is generated by in vitro transcription of a corresponding RNA construct and quantified by measurement of the optical density at 260 nm (OD 260 nm). A defined number of these molecules is mixed with a specified amount of total RNA from naive Huh-7 cells and this mixture is transfected into naive Huh-7 cells with the aid of electroporation. At the same time the total RNA of a cell clone, produced by the method described in Figure 1, is isolated using a known state of the art method and the amount of HCV-RNA contained therein is determined by means of Northern Blot using a HCV specific RNA probe and subsequent quantification via phosphoimaging. A defined amount of this total RNA is analogously transfected in naïve Huh-7 cells. These cells from both the cultures are then subjected to a G418 selection and the number of colonies created is determined by counting after fixing and staining with Coomassie Brilliant Blue. For the determination of transfection efficiency 1 $\mu$ g of a

plasmid allowing the expression of luciferase is added to each transfection culture. An aliquot of the transfected cells is collected after 24 hours and the luciferase activity determined in the  
5 respective cell lysates. The number of colonies is always normed to the luciferase expression.

**Fig. 10:** Sequence analysis of the 9-13 clone. Total RNA of the 9-13 cell clone, resulting from transfection of HCV-RNA construct I377/NS3-3', was isolated using a known state of the art method and the HCV-RNA construct amplified from nucleotide position 59 to 9386 with the aid of 'long-distance RT-PCR' using primer S59 and A9413. The  
10 PCR fragments were cloned and 11 clones (9-13 A - K) completely sequenced. Clones D and I, E and G as well as H and J turned out to be identical, respectively. The positions of the amino acid  
15 differences in the NS3-5B region between the recloned HCV-RNAs and the parental construct are marked with a thick vertical mark for each clone. Each clone was digested with restriction enzyme  
20 SfiI and the respective fragment inserted in the parental construct. These clones were each transfecte  
25 in Huh-7 cells and the cells subjected to selection as described in Figure 1. The number of cell clones obtained with each construct is noted next to the respective construct on the right.  
30

**Fig.11 A:** Principle of determination of replication with the aid of a reporter gene. In the upper part of the figure, the HCV-DNA construct I<sub>389</sub>/Luc/NS3-3' is shown. It consists of the HCV 5' NTR (nucleotide positions 1-389), the luciferase gene (luc), the IRES of the encephalomyocarditis virus, the HCV NS3-5B and the 3' NTR. The  
35

position of the active centre of the NS5B RNA polymerase, into which a deactivating amino acid exchange was introduced, is indicated by 'GND'.  
5 The plasmids, which code for the HCV-RNA construct which is able to replicate or is defective, are digested with the restriction enzyme ScaI and added to an in vitro transcription reaction with T7 RNA polymerase.  
10 After removal of the template DNA, the respective HCV-RNA constructs were transfected in naive Huh-7 cells by means of electroporation and the latter collected at regular intervals.

15 **Fig.11 B:** Comparison of luciferase activity in cells transfected with the parental HCV-RNA construct I<sub>389</sub>/Luc/NS3-3'/wt (wt) or the following variants:  
inactive RNA (318 DN), variants 9-13F or variant  
5.1. The cells were collected 6 (not shown), 24,  
48, 72, 96, 120, 144 and 168 hours after  
20 transfection and luciferase activities determined  
by luminometric measurement.

25 **Fig. 12:** Selectable HCV full-length genomes (constructs I<sub>389</sub>/core 3'/5.1 and I<sub>389</sub>/core 3'/9-13F).  
(A) Diagram of the full-length construct. The region between both indicated recognition sites for the restriction enzyme SfiI corresponds to the sequences of the highly adapted RNA variants 5.1. or 9-13F.  
30 (B) Number of colonies which were obtained after transfection of 0.1 µg in vitro transcribed RNA of the construct I<sub>389</sub>/core-3'/5.1. described under A into HUH7-cells. The result of a representative experiment is given.  
35 (C) Demonstration of autonomously replicating HCV full-length RNAs in G418 resistant cell clones which were obtained after transfection of the

corresponding in vitro transcript. The illustration shows the autoradiogram of a Northern Blot, which was hybridised with a probe against the neo-resistance gene and the HCV 5' NTR. The controls shown in lanes 1 and 2 each correspond to  $10^6$  molecules of the indicated in vitro transcripts, mixed with total RNA from naive Huh-7 cells. The negative control contains only total RNA from naive Huh-7 cells (lane 3). Lanes 4-9 contain 3-10 µg total RNA from G418 resistant cell clones, which were obtained after transfection by in vitro transcribed I<sub>389</sub>/core 3'/5.1 RNA or I<sub>389</sub>/core 3'/9-13F RNA. The G418 concentration used for the selection is indicated in each case. Five of the cell clones shown contain the highly adapted RNA variant 5.1 (lanes 4-8), one contains the adapted RNA variant 9-13F (lane 9).

Fig. 13: HCV-RNA constructs with a reporter gene. (A) Bicistronic HCV-RNA constructs. The reporter gene is translated with the aid of a separate IRES. (B) Monocistronic HCV-RNA constructs. The reporter gene product is expressed as fusion protein with a HCV protein. Both portions are linked via a recognition sequence for a viral or cellular protease, which permits a proteolytic separation of the two fused protein portions. In the example shown the reporter gene product and respective HCV protein was fused through a recognition sequence for ubiquitin (Ub).

Fig. 14: Tricistronic full-length HCV-RNA construct, that in addition to the resistance gene possesses an inserted foreign gene (ther. gene).

Fig. 15: Monocistronic HCV-RNA constructs, for which the resistance gene product is expressed as a fusion protein with HCV portion. The resistance gene (RG) is either active as a fusion protein or it is fused with the HCV portion via a proteolytically cleavable sequence in such a way that the resistance gene product is split from the HCV portion by a cellular or viral protease. In the example shown the resistance gene was fused with the respective HCV portion through the sequence coding for ubiquitin (Ub).

15 Example 1: Production of HCV-RNA constructs

(A) Synthesis and cloning of a complete HCV consensus genome by means of RT-PCR

20 The HCV genome, which means the HCV-RNA, was isolated from the liver of a chronically infected patient as described in the following:

The total RNA was isolated from approx. 100 mg of liver  
25 according to the method described by Chomczynski and Sacchi (1987, Anal. Biochem. 162, 156). Using 1 µg of this isolated RNA a reverse transcription was carried out with the primers A6103 (GCTATCAGCCGGTTCATCCACTGC) or A9413 (CAGGATGGCCTATTGG CCTGGAG) and the expand reverse  
30 transcriptase system (Boehringer Mannheim, Germany) according to the manufacturer's recommendations. A polymerase chain reaction (PCR) was carried out with the products of this reverse transcription using the expand long template system (Boehringer Mannheim, Germany), in  
35 which the buffer containing 2 % of dimethyl sulfoxide was used. After one hour at 42 °C, 1/8 of the reaction mixture was used for a first PCR with primers A6103 and S59

(TGTCTTCACGCAGAAAGCGTCTAG) or A9413 and S4542 (GATGAGCT CGCCCGCGAAGCTGTCC). After 40 cycles, 1/10 of this reaction mixture was used for a second PCR with primers S59 and A4919 (AGCACAGCCCCGCGTCATAGCACTCG) or S4542 and A9386 5 (TTAGCTCCCCG TTCATCGGTTGG. After 30 cycles the PCR products were purified by means of preparative agarose gel electrophoresis and the eluted fragments were ligated into the vector pCR2.1 (Invitrogen) or pBSK II (Stratagene). Four clones of each fragment were analyzed and sequenced, 10 and a consensus sequence was established. For this purpose the DNA sequences were compared to each other. The positions, where the sequence of one fragment was different from the others, were considered as undesired mutations. In the case of ambiguities of the sequence, shorter 15 overlapping PCR fragments of the respective region were amplified and several clones sequenced. By this means several potential mutations could be identified in each fragment and consequently an isolate specific consensus sequence could be established. This established consensus 20 sequence or genome belongs to the worldwide spread genotype 1b. The non-translated region at the 3' end (=3' NTR) was obtained by means of a conventional PCR, whereby an antisense primer, which covers the last 24 nucleotides of the 'X-tail' known from the state of the art was used 25 (Tanaka et al., 1995, Biochém. Biophys. Res. Commun. 215, 744; und Rice, PCT/US 96/14033). The authentic non-translated region on the 5' end (=5' NTR) downstream of the T7 promoter was created by means of PCR, whereby an oligonucleotide, which corresponds to a shortened T7 30 promoter (TAA TAC GAC TCA CTA TAG) and the first 88 nucleotides of HCV, was used on one hand and one of the previously mentioned plasmids, which carries one of the 5' fragments of the genome, was used on the other hand. A complete HCV consensus genome was assembled from the 35 subgenomic fragments with the smallest number of non-consensus replacements and inserted into a modified pBR322 vector. Deviations from the consensus sequence were

eliminated by means of site-directed mutagenesis. In order to produce run-off transcripts with an authentic 3' end, the 3' NTR of the isolates (with the end TGT) was modified to AGT (according to the sequence of the genotype 3 = clone 5 'WS' according to Kolykhalov et al., 1996, J. Virol. 70, 3363) and an additional nucleotide replacement was carried out at position 9562, to preserve the A:T pairing in the hairpin structure at the 3' end of the 3' NTR (Kolykhalov et al. ibid.). In order to eliminate an internal restriction 10 site for the Scal enzyme, a silent nucleotide replacement was further carried out. After joining the full-length genome with the appropriate 5' and 3' NTRs the complete HCV sequence was analyzed. No undesired nucleotide replacement was detected.

15

The HCV genome produced in this way should be hepatotropic according to the definition.

20

(B) Synthesis of selectable HCV subgenomic constructs

By means of the consensus genome described under (A), HCV subgenomic constructs were created, which include the antibiotic resistance gene neomycin phosphotransferase (NPT) and two sequences of internal ribosome entry sites (IRES). The biochemical procedures used for this are known and familiar to the expert (see: Sambrook, J., E.F. Fritsch, T. Maniatis, 1989, Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbour Laboratory, Cold Spring Harbor, N.Y.; Ausubel et al. (eds.), 1994, Current Protocols in Molecular Biology, Vol. 1-3, John Wiley & Sons Inc., New York). The antibiotic resistance gene was inserted immediately after the 5' NTR, obtaining a bicistronic RNA as a result (see Fig. 1 A). However, the antibiotic resistance gene might as well be inserted at another site of the HCV subgenomic construct, for example within the nucleotide sequence encoding for the polyprotein, obtaining a monocistronic RNA as a result (see

Fig. 5 A), or in the 3' NTR (see Fig. 5 B). The IRES elements are one of the two HCV-IRES variants from nucleotides 1-377 or nucleotides 1-389 as well as the IRES of the encephalomyocarditis virus, which controls the 5 translation of the HCV sequence downstream from the genes for NS2 or NS3 up to the authentic 3' end of the genome.

The two HCV-IRES variants mentioned were determined in the following way:

10 Based on deletion analyses of the 3' end of the HCV-IRES (Reynolds et al. 1995, *EMBO J.* 14, 6010) various segments of the 5' NTR were fused with the NPT gene and analyzed with regard to the maximum number of colonies formed by means of 15 co-transfection with a plasmid containing the T7 RNA polymerase gene. The best results were achieved with the HCV sequences from 1-377 to 1-389. Due to the fact that the AUG start codon of the HCV polyprotein is located at position 342 and is consequently included in the IRES 20 sequence, 12 or 16 amino acids of the HCV core protein ("core protein") were fused with the neomycin phosphotransferase (see Fig. 1 A).

Accordingly, these modified HCV subgenomic constructs were 25 given the designations  $I_{377}/NS2-3'$  (or  $I_{377}/NS3-3'$ ) and  $I_{389}/NS2-3'$  (or  $I_{389}/NS3-3'$ ). They are illustrated in a diagram in Fig. 1A.

Different cell lines and primary cell cultures of human 30 hepatocytes were transfected with *in vitro* transcripts of these modified parental HCV subgenomic constructs  $I_{377}/NS2-3'$  (or  $I_{377}/NS3-3'$ ) and  $I_{389}/NS2-3'$  (or  $I_{389}/NS3-3'$ ).

As a parallel negative control for all transfection 35 experiments, an appropriately modified, but defective subgenome was constructed for every modified parental HCV subgenomic construct, which differs from the parental

construct due to the fact, that it has a deletion of 10 amino acids within the reading frame comprising the active site of the NS5B RNA polymerase (Behrens et al., 1996, *EMBO J.* 15, 12; and Lohmann et al., 1997, *J. Virol.* 71, 8416).

5

(C) Synthesis of selectable HCV genomic constructs

A NS2-3' subgenomic construct, linked at its 5' end to a fragment of the luciferase gene and the complete EMCV-IRES, 10 was restricted with NcoI and SpeI and purified using preparative agarose gel electrophoresis. The vector obtained in this way was ligated with a NcoI/NotI-HCV fragment corresponding to the nucleotide positions 342 to 15 1968 of the HCV genome and a NotI/SpeI fragment corresponding to the nucleotide positions 1968-9605 in a 3-factor-ligation. The resulting construct, where the complete HCV open reading frame and the 3' NTR lie downstream from the luciferase gene fragment and the EMCV-IRES, was then restricted with PmeI and SpeI and ligated 20 with the analogously restricted I<sub>389</sub>/NS3-3'/wt subgenomic construct vector. This selectable HCV genomic construct is illustrated in Fig. 4.

25 (D) Production of in vitro transcripts corresponding to the HCV-RNA constructs

The previously described purified plasmid DNAs were linearized with ScaI and used for an in vitro transcription reaction after phenol/chloroform extraction and isopropanol 30 precipitation using the following components: 80 mM HEPES, pH 7.5, 12,5 mM MgCl<sub>2</sub>, 2 mM Spermidine, 40 mM Dithiothreitol, 2 mM of each NTP, 1 unit RNasin/μl, 50 μg/ml restricted DNA and approx. 2 units/μl T7 RNA polymerase. After 2 hrs. at 37 °C half of the amount of T7 35 polymerase was added and the reaction mixture was incubated for two further hours. In order to remove DNA the mixture was extracted with acid phenole (U. Kedzierski, J.C. Porte,

1991, Bio Techniques 10, 210), precipitated with isopropanol, the pellet was dissolved in water and incubated with DNase (2 units per  $\mu$ g DNA) for 60 min. at 37 °C. After subsequent extraction with acid phenole, acid phenole/chloroform and chloroform as well as isopropanol precipitation the dissolved RNA was quantified using optical density measurement and its integrity was checked using formaldehyde-agarose gel electrophoresis.

10

Example 2: Transfection experiments with the hepatoma cell line Huh 7

With all transfection experiments it was carefully ensured that any template DNA had been removed beforehand so as to avoid the possibility of this DNA integrating in transfected cells and conferring a neomycin resistance to them independent from HCV replication. The reaction mixture was therefore treated with 2 units of DNase per  $\mu$ g DNA for 60 min. at 37 °C after the in vitro transcription and extracted with acid phenole, acid phenole/chloroform and chloroform. Prior to being used for the transfection the precipitated RNA was analyzed using formaldehyde-agarose gel electrophoresis.

25

Three separate transfection experiments were carried out with the highly differentiated human hepatoma cell line Huh 7 (according to Nakabayashi et al. 1982, Cancer Res. 42, 3858). Each time 15  $\mu$ g RNA were introduced into  $8 \times 10^6$  Huh 7 cells by electroporation and the cells were seeded in culture dishes with a diameter of 10 cm. 24 hours after seeding, neomycin (= G418) was added in a final concentration of 1 mg/ml. The culture medium was changed twice a week. After 3 - 5 weeks small colonies were visible, which were isolated and grown under the same culture conditions.

The cell clones obtained during the first experiment were isolated and subpassaged. Most of the clones died during this procedure, and the final yield was only 9 clones from cells, which had been transfected with the parental HCV 5 subgenomic constructs and 1 clone (clone 8-1) from cells, which had been transfected with a defective HCV genomic construct, namely a defective NS2-3' HCV-RNA. Apart from an extended doubling time and the occasional occurrence of irregularly shaped cells, no consistent morphological 10 differences were found between these 9 cell clones and the single cell clone (clone 8-1) or the parental Huh 7 cells.

The main criteria for functioning HCV genomic constructs are the formation of viral RNA with the correct size and 15 the absence of (integrated) plasmid DNA, which could transfer or mediate on a G418 resistance.

To determine the HCV-RNA in the Huh 7 cells, the total RNA was isolated and analyzed by means of the common Northern 20 Blot method using a plus-strand specific ribo probe (RNA probe). For this purpose the total RNA was isolated from the respective cell clones according to the method described by Chomczynski and Sacchi 1987, *Anal. Biochem.* 162, 156 and 10 µg RNA, which is equivalent to the total 25 RNA content of 0,5 - 1 x 10<sup>6</sup> cells, are separated using denatured formaldehyde-agarose gel electrophoresis (lanes 3 to 12 of Fig. 1 B). At the same time 10<sup>9</sup> in vitro transcripts (ivtr.), which correspond to the I<sub>389</sub>/NS2-3'/wt or the I<sub>389</sub>/NS3-3'/wt replicon RNAs, are separated as well 30 as size markers with authentic sequence (lane 1 or lane 2). The separated RNA was transferred onto nylon-based membranes and hybridized with a radioactively labeled plus-strand specific RNA probe, which was complementary to the complete NPT gene and the HCV-IRES of nucleotide 377 to 35 nucleotide 1. The positions of the HCV specific RNAs (arrows) and the 28S rRNA are specified to the right of lane 12, the size (amount of nucleotides) of the RNA marker

is specified to left of lane 1. The RNA marker fragments contain HCV sequences and therefore hybridize with the ribo probe (= RNA probe). The results of this analysis are illustrated in Fig. 1 B.

5

With the exception of clone 8-1 transfected with the defective HCV genomic construct, all cell clones produced homogenous HCV-RNAs of correct lengths (approx. 8640 nucleotides in the case of the NS2-3' and approx. 7970 nucleotides in the case of NS3-3' replicon). This result is an indication of the fact, that the functional replicons or the functional HCV genomic constructs transfer the G418 resistance. In order to exclude, that the G418 resistance is caused by a plasmid DNA, which is integrated in the genome of the Huh 7 host cell and transcribed under control of a cellular promoter, the DNA of each clone was analyzed by means of a NPT gene specific PCR. Consequently the DNA was isolated from the selected Huh 7 cell clones by means of digestion with proteinase K (40 $\mu$ g/ml, 1h, 37 °C) in 10mMTris, pH7,5, 1mM EDTA, 0,5 % SDS and subsequent extraction with phenol, phenol/chloroform and isopropanol precipitation. The DNA precipitate was dissolved in 10 mM Tris (pH 7,5) and 1 mM EDTA and incubated with Rnase A for 1 hour. Following a phenol/chloroform extraction and ethanol precipitation, 1  $\mu$ g DNA, equivalent to 4 - 8  $\times$  10<sup>4</sup> cells, was analyzed by means of PCR using NPT gene specific primers (5'-TCAAGACCGACCTG TCCGGTGCCC-3' and 5'-CTTGAGCCTGGCGAACAGTTCGGC-3'), and a DNA fragment consisting of 379 nucleotides was generated. The specificity of the PCR product was established by means of the Southern Blot method, in which a DNA fragment labeled by digoxigenin was used, which corresponds to the NPT gene. As positive controls (for the detection of possibly contaminating nucleic acids) the PCR method was carried out with 10<sup>7</sup> plasmid molecules or 1  $\mu$ g DNA from a BHK cell line, which was stably transfected with a neomycin resistance gene, and

as negative control, the PCR was carried out with the same reagents but without added DNA.

The results of this analysis are illustrated in Fig. 1 C.

5 The lanes 1 and 2 represent the positive controls, lane 13 represents the negative control. The numbers to the left of lane 1 indicate the sizes of the nucleotide marker molecules.

10 A NPT-DNA could not be detected in any cell clone, apart from clone 7-3 (Fig.1C, lane 3), which was obtained from cells after transfection with a NS2-3' replicon/NS2-3'HCV genomic construct, and clone 8-1 (Fig.1C, lane 12), which was obtained from cells after transfection with a defective 15 HCV genomic construct. This result was another indication of the fact, that the G418 resistance of most clones was passed on by the replicating HCV-RNA. But even regardless of these results, it is unlikely, that integrated plasmid DNA produces HCV-RNAs of correct sizes, because the 20 plasmids used for the in vitro transcription contain neither an eukaryotic promoter nor a polyadenylation signal. In the case of clone 7-3 the resistance is therefore very probably passed on by the HCV-RNA construct or the replicating HCV-RNA as well as by an integrated NPT 25 DNA sequence, whereas the resistance of the cells of clone 8-1 is only caused by the integrated plasmid DNA.

Clone 9-13 (Fig. 1 B, lane 11) was subject to further tests, to confirm that the G418 resistance was passed on by 30 a self-replicating HCV-RNA. Clone 8-1, which carries the integrated copies of the NPT gene, was used throughout as negative control. A PCR was carried out, which allows the detection of < 1000 NPT gene copies in ~ 40.000 cells, with the aim to rigorously exclude the presence of NPT-DNA in 35 clone 9-13. The result of this PCR is illustrated in Fig. 2A. The PCR proceeded in detail as follows:

During the test,  $10^6$  -  $10^2$  plasmid molecules ( $I_{377}/NS3-3'$ /wt) were used either directly (lanes 7 - 11) or after adding each 1  $\mu$ g 9-13 DNA (lanes 2 - 6). The specificity of the amplified DNA fragment was determined by Southern Blot using a NPT specific probe. A PCR without DNA probe was carried out as negative control (lane 12).

Even with this sensitive method, no plasmid DNA could be detected in one  $\mu$ g DNA of the cell clone 9-13 (lane 1). To estimate the amount of HCV plus- and minus-strand RNAs in these cells, a dilution series of total RNA was analyzed with the Northern Blot method using a plus- or minus-strand specific radioactively labeled ribo probe (= RNA probe). For this purpose, 8, 4 or 2  $\mu$ g of total RNA, which have been isolated from the cell clones 9-13 and 8-1, were analyzed in parallel to known amounts of in vitro transcripts with plus- or minus-strand polarity (control RNAs) in the Northern Blot method and were then subjected to a hybridization. The hybridization was carried out with a plus-strand specific ribo probe, which covered the complete NPT gene and the HCV-IRES ('plus-strand', top panel), or with a minus-strand specific RNA probe, which was complementary to the NS3 sequence ('minus-strand', bottom panel). The arrows mark the positions of replicon RNA. The results of this analysis are illustrated in Fig. 2 B.

Approx.  $10^8$  copies/ $\mu$ g total RNA were detected in the case of the plus-strand, which is equivalent to 1000 - 5000 HCV-RNA molecules per cell, whereas the amount of minus-strand RNA was 5- to 10-times less. This result corresponds to the assumption, that the minus-strand RNA is the replicative intermediate form or intermediate copy, which is used as a template for the synthesis of the plus-strand molecules.

35

Due to the fact that the reaction is mainly catalyzed by the viral RNA-dependent RNA polymerase, the synthesis of

the HCV-RNAs should be resistant to dactinomycin, an antibiotic, which selectively inhibits the RNA synthesis of DNA templates, but not the RNA synthesis from RNA templates. To confirm this assumption, cells were incubated  
5 with [<sup>3</sup>H] uridine in the presence of dactinomycin, the radioactively labeled RNAs were extracted, separated by means of denaturing agarose gel electrophoresis and analyzed with the aid of a common Bio-Imager using a [<sup>3</sup>H]-sensitive screen. For this purpose approx. 5 x 10<sup>5</sup> cells of  
10 the clones 9-13 and 8-1 were incubated at a time with 100 µCi [<sup>3</sup>H] uridine for 16 hrs. in the absence (-) or presence (+) of 4 µg/ml of dactinomycin (Dact). Following this labeling reaction the total RNA was prepared and analyzed by means of formaldehyde-agarose gel electrophoresis. Only  
15 1/10 of the total RNA is illustrated in the first two lanes. The radioactively labeled RNA was visualized using a BAS-2500 Bio-Imager (Fuji).

The results of this analysis are illustrated in Fig. 2 C.  
20 In accordance with the inhibitor profile of the NS5B polymerase (Behrens et al., 1996, *EMBOJ.* 15, 12 and Lohmann et al., 1997, *J. Virol.* 71, 8416), dactinomycin had no influence on the replication of the HCV RNA, whereas the synthesis of cellular RNA was inhibited. A RT-PCR was  
25 carried out for recloning the replicating sequences, to confirm the identity of the viral RNA. The sequence analysis of the recloned RNA showed that the RNA in clone 9-13 is HCV specific and corresponds to the transfected transcript of the HCV construct I<sub>377</sub>/NS3-3'/wt.

30 For the analysis of the viral proteins, at first the respective cells were metabolically radioactively labeled with [<sup>35</sup>S] methionine/cysteine, then lysed and afterwards the HCV specific proteins were isolated from the cell  
35 lysates by means of immunoprecipitation. The results of these analyses are illustrated in Fig. 3 A. The detailed procedure was as follows: Cells of the cell clones 9-13

(wt) and 8-1 ( $\Delta$ ) were metabolically radioactively labeled by treating them with a protein labeling mixture familiar to the expert and available on the market (for example. NEN Life Science). The HCV specific proteins were separated 5 from the cell lysate by immunoprecipitation (IP) under non-denaturing conditions (for example according to Bartenschlager et al., 1995, J. Virol. 69, 7519) using three different antisera (3/4, 5A, 5B, according to the labeling on the top end of the lanes 1 to 12). The immune 10 complexes were analyzed by means of tricine SDS-PAGE and made visible by means of autoradiography. To obtain authentic size markers, the homologous replicon construct I<sub>377</sub>/NS3-3'/wt was subject to a transient expression by the vaccinia virus-T7 hybrid system in the Huh 7 cells. The 15 resulting products were used as size markers (lanes 7 - 9) parallel to the cells of the clones 9-13 and 8-1. Identified HCV proteins are labeled on the left edge of lane 1, the molecular weights (in Kilodalton) are specified on the right edge of lane 9. It should be noted that the 20 NS3/4 specific antiserum ('3/4') used preferably reacts with NS4A and NS4B causing an underrepresentation of NS3.

All viral antigens could unambiguously be detected, and their apparent molecular weights did not show any 25 difference to those being detected after transient expression of the same bicistronic HCV-RNA construct in the original Huh 7 cells. An immunofluorescence detection reaction was carried out using NS3 and NS5A specific antisera, to determine the subcellular distribution of the 30 viral antigens (for example according to Bartenschlager et al., 1995, J. Virol. 69, 7519). For this purpose cells of the clones 9-13 (wt) and 8-1 ( $\Delta$ ) were fixed with methanol/acetone 24 hrs. after incubating on coverslips and incubated with polyclonal NS3 or NS5A specific antisera. 35 The bound antibodies were made visible with a commercially available FITC conjugated anti-rabbit antiserum. The cells

were counterstained with the 'Evans Blue' stain to suppress unspecific fluorescence signals.

The results of this detection test are illustrated in Fig.  
5 3 B. A strong fluorescence in the cytoplasm could be detected with both antisera. The NS5A specific antiserum also caused a slight nuclear fluorescence, which indicates that at least small amounts of this antigen also reach the nucleus. But the generally dominating presence of the viral  
10 antigens in the cytoplasm are a strong indication that HCV-RNA replication occurs in the cytoplasm, as is the case with most RNA viruses.

These results prove clearly that the establishment of a  
15 cell culture system for the HCV could be accomplished with the test arrangement described, the efficiency of which surpasses everything known up until now by far and for the first time allows the detection of viral nucleic acids and proteins with conventional and approved biochemical  
20 methods. This efficiency actually allows detailed examination of HCV pathogenesis, genetic analyses of different HCV functions and a precise study of the virus/host cell interaction, through which new starting points for the development of a antiviral therapy can be  
25 defined.

Example 3: Transfection of Huh 7 cells with HCV total constructs

30 Huh7 cells are transfected and selected as described in example 2, whereby in this case selectable constructs are used, which contain the complete virus genome. Corresponding to Example 2, the resulting cell clones are  
35 tested for the absence of HCV-DNA by means of PCR and the productive replication of HCV-RNA is then established by means of Northern Blot, [<sup>3</sup>H]uridine labeling in the

presence of dactinomycin, detection of the viral proteins or antigens preferably with the aid of the Western Blot, the immunoprecipitation or immunofluorescence. In contrast to the arrangements described in Example 2, the construct 5 described here makes it possible to obtain more complete and very likely infectious viruses, which has not been the case in the subgenomic constructs described in Example 2. These viruses existing in the cell and the cell culture supernatant are concentrated for example by means of 10 ultracentrifugation, immunoprecipitation or precipitation with polyethyleneglycol, and all exogenous nucleic acids, which means those that are not incorporated into the virus particle, are digested by incubation with nucleases (RNase, DNase, micrococcal nuclease). In this way all contaminating 15 nucleic acids, which are not included in the protecting virus particle, can be removed. After inactivation of the nucleases, the protected viral RNA is isolated for example by means of incubation with proteinase K in a buffer containing SDS, by extracting with phenol and 20 phenol/chloroform and detected by means of Northern Blot or RT-PCR using HCV specific primers. Also in this test arrangement, the combination of the HCV consensus genome described with a selection marker was crucial for the efficient production of viral RNA, viral protein and 25 therefore HCV particles.

Example 4: Production and application of a HCV-  
RNA construct, whereby the resistance gene is  
linked to the HCV subgenomic sequence via a  
ribozyme or a recognition site for a ribozyme.

A HCV-RNA construct is produced according to Example 1 or Example 3, where an antibiotic resistance gene is linked to 30 the HCV-RNA sequence through a ribozyme or a recognition sequence for a ribozyme. These constructs are illustrated in a diagram in Fig. 7. Huh 7 cells are transfected with 35

this HCV-RNA construct as described in Example 2. A selection with the appropriate antibiotic follows the transfection into the cells. The inserted ribozyme is activated in the cell clones obtained in the procedure or,  
5 in the case of a construct, which carries a recognition sequence for a ribozyme, the ribozyme is transfected into the cell (for example by means of transfection of a ribozyme construct or infection with a viral expression vector, into which the respective ribozyme has been  
10 inserted). In both cases the resistance gene is separated from the HCV-RNA sequence by the ribozyme-dependent cleavage. The result in the case of the HCV genome is an authentic HCV genome without a resistance gene, which can form authentic infectious virus particles. A HCV replicon  
15 without resistance gene is created in the case of the HCV subgenomic constructs.

**Example 5: Co-transfection of a HCV-RNA construct with a**  
20 **separate luciferase transfection construct**

A HCV-RNA construct is produced according to Example 1 (A) or Example 3 or Example 4. At the same time a transfection construct is produced, which comprises the luciferase gene,  
25 whereby this luciferase gene is linked to a first nucleotide sequence, which encodes a HCV protease (for example NS3 protease) cleavage site, to a second nucleotide sequence, which encodes for another protein or a part of another protein. HCV-RNA construct and transfection  
30 construct are transfected into any host cells, preferably hepatoma cells, most preferably Huh 7 cells. This can be realised as described in Example 2. The product of the modified luciferase gene is a luciferase fusion protein, where the luciferase is inactivated due to the fusion with  
35 the foreign part. The fusion protein, which contains a recognition sequence for a HCV protease, is cleaved in transfected cells with high HCV replication, and

consequently the active form of the luciferase, which can be identified through luminometric measurement, is released. If the replication of the HCV-RNA construct is inhibited the fusion protein will not be cleaved and no active luciferase will be released. The quantitative determination of the luciferase is therefore a measure for the replication of the HCV construct. Instead of the luciferase gene, another reporter gene can just as easily be used, which is modified in the same way, so that its expression depends on the viral replication, although this reporter gene is not part of the HCV construct. A cellular protein, which is deactivated or activated by the HCV proteins or nucleic acid, can also be used as a so called surrogate marker. The expression or activity of this surrogate marker is in this case a measure for the replication of the viral DNA.

Example 6: Production of HCV subgenomic constructs with integrated foreign genes to be used as liver cell specific vector in gene therapy

These recombinant and selectable HCV subgenomic constructs are transfected in transcomplementing helper cell lines, which means in cell lines, which express the missing functions (for example the structural proteins) in an inducible or constitutive way. Cell clones containing a functional HCV subgenomic construct can be established through appropriate selection. The viral structural proteins expressed from the host cell allow the formation of virus particles, into which the RNA of the HCV subgenomic constructs will be transfected. The results are therefore virus-like particles, which contain a HCV subgenomic construct according to the invention including the inserted foreign gene and which can transmit this to other cells by means of infection. An example for this construct is illustrated in Fig. 8.

It is also possible to use the HCV subgenomic construct with integrated foreign gene directly as an expression vector. This involves the same method as mentioned 5 previously except that cell lines, which do not express transcomplementing factors, are transfected. In this case the HCV construct is only used as an expression vector.

10 **Example 7: Production of cell culture adapted HCV-RNA constructs**

(A) **Method of isolation**

15 The following method was used to determine adaptive mutations and to produce cell culture adapted HCV-RNA constructs: cells were transfected with a HCV-RNA construct as described in Examples 1 and 2 and G418-resistant cell clone produced. For the determination of ability to 20 replicate (understood in this context to be the number of G418 resistant cell clones obtained per microgram of transfected HCV-RNA or HCV-RNA construct), the total RNA from one of the cell clones, [the 9-13 clone (Fig. 1 B, lane 11)], was isolated and the quantity of HCV-RNA 25 contained within it was determined by Northern Blot as described in Fig. 2 B. Ten micrograms of the total RNA, containing approx.  $10^9$  molecules of HCV-RNA, was then transfected into naive Huh-7 cells using electroporation (Fig. 9). In parallel,  $10^9$  in vitro transcripts of the 30 analogous neo-HCV-RNA, which had been adjusted with isolated total RNA from naive Huh-7 cells to a total RNA quantity of 10  $\mu$ g, were transfected in naive Huh-7 cells. After selection with G418, the number of cell colonies, expressed as 'colony forming units (cfu) per microgram 35 RNA', was determined in both these cultures. At a concentration of 500  $\mu$ g/ml G418 in the selection medium, the number of colonies obtained with the HCV-RNA contained

in **isolated** total RNA from clone 9-13, was approx. 100,000 cfu per microgram HCV-RNA. In contrast, only 30 - 50 colonies were obtained with the same quantity of in vitro transcribed HCV-RNA. This result confirms that the specific 5 infectivity of the HCV-RNA isolated from the cell clones is approx. 1,000 - 10,000 times higher than the infectivity of the analogous in vitro transcripts. The experimental approach is shown in Fig. 9.

10 With the aid of 'long-distance RT-PCR', the HCV-RNA was amplified from the total RNA of the 9-13 cells, the PCR amplificate was cloned and numerous clones were sequenced. A comparison of the sequences of these recloned RNAs with the sequence of the RNA originally transfected into the 15 naive Huh-7 cells, showed that the recloned RNAs possessed numerous amino acid exchanges distributed over the whole HCV sequence (Fig. 10). *SfiI* fragments of these recloned mutants were used to replace the analogous *SfiI* fragment of the original replicon construct, and RNAs of the 20 respective mutants were transfected in naive Huh-7 cells. After selection with G418 the number of colonies created was determined for each HCV-RNA mutant. While only 30 - 50 colonies per microgram RNA were obtained with the parental RNA the number of colonies was noticeably higher for two of 25 the recloned variants (Fig. 10). In the case of the HCV-RNA constructs 9-13I and 9-13C the specific infectivity was increased to 100 - 1,000 cfu per microgram RNA and for 9-13F replicon it was 1,000 - 10,000 cfu per microgram RNA. These results show that the amino acid exchanges in the 30 analysed NS3-5B regions of the mutants 9-13I, 9-13C and particularly of 9-13F, led to a considerable increase in ability to replicate. In contrast all the other HCV-RNA constructs (9-13 A, B, G, H und K) were no longer able to replicate, they thus contained lethal mutations.

35

In order to answer the question which of the amino acid exchanges in the 9-13F construct led to an increase in

replication, the exchanges were introduced separately or in combination into the parental HCV-RNA construct, and the corresponding RNAs transfected in naive Huh-7 cells. The result of the transfection with these RNAs is summarised in  
5 Table 1. From this it is evident that in the present example the high ability to replicate is determined by several mutations. The amino acid exchanges in the HCV-RNA regions NS5A and NS4B make the greatest contribution. The single exchanges in the NS3-Region also make a contribution  
10 and perhaps they are synergistic.

These results confirm that it was through the G418 selection of the cells transfected with the neo-HCV-RNA construct that there was enrichment of those HCV-RNAs  
15 having noticeably higher ability to replicate. HCV-RNA constructs with greatly differing replication efficiencies can be selected using the experimental approach described here. The higher the concentration of the antibiotic in the selection medium, in/on which the HCV-RNA construct  
20 containing cells are cultivated for selection, the higher must be the extent of adaptive mutations and hence replication efficiency of the relevant HCV-RNA constructs, to allow the cells to grow under these conditions. If the  
25 selections are carried out using lower antibiotic concentrations, cells can survive and multiply, but the HCV-RNA construct shows a comparatively lower replication efficiency and fewer adaptive mutations.

As has been shown, the 9-13F HCV-RNA construct described so  
30 far, which contains several adaptive mutations, had a higher replication efficiency than the parental HCV-RNA. In order to obtain HCV-RNAs with even higher replication in cell culture, the HCV-RNA contained in the total RNA of a selected cell clone was passaged several times in naive  
35 Huh-7 cells. The selected 5-15 cell clone, was obtained by transfection with the HCV-RNA construct I<sub>389</sub>/NS3-3' (Fig. 1). It largely corresponds to the cell clone 9-13, produced

by transfection with a HCV-RNA construct, having a HCV-IRES shorter by 22 nucleotides ( $I_{389}/NS3-3'$ ; Fig. 1). Ten micrograms of total RNA, isolated from cell clone 5-15, were transfected into naive Huh-7 cells using electroporation and the cells subjected to a selection with 1 mg/ml G418. The total RNA from one of the cell clones thus produced was again isolated, transfected into naive Huh-7 cells and selected in the same way. This process was repeated a total of four times. After the fourth passage the total RNA was isolated from a cell clone and the neo-HCV-RNA amplified with the aid of the 'long-distance RT-PCR'. The amplified DNA fragment was digested with the restriction enzyme *SfiI* and inserted into the *SfiI*-restricted parental construct  $I_{389}/NS3-3'$ . Over 100 DNA clones were obtained altogether and then analysed by means of restriction digestion. In vitro transcribed RNA of about 80 of these clones was each transfected into naive Huh-7 cells and subjected to a selection with 500mg/ml G418. Of the 80 neo-HCV-RNA variants examined, the great majority proved to be replication defective. However, the specific infectivity, expressed as 'colony forming units' per microgram RNA, was noticeably increased in the case of two mutants, 5.1 and 19 (Table 2). Through several passages of the RNA in cell culture it is clear that HCV-RNAs are produced whose replication efficiency due to mutations ("adaptive mutations") is several orders of magnitude higher than the original RNA cloned from patients.

#### (B) Modified method

Adaptive mutations produced and identified according to (A) can be transferred into a HCV-RNA construct with low ability to replicate. This leads to a huge increase in the replication of this construct. The increase is so great it can be demonstrated that HCV-RNAs transfected into cell culture can replicate even in the absence of selection pressure Fig. 12 shows a comparison of the replication

efficiency of HCV-RNAs, which corresponded either to the starting sequence or to the adaptive sequences 9-13F or 5.1. For the purposes of simple measurement, the neo-gene was removed and replaced by the gene for luciferase. The 5 negative control used was again a HCV-RNA construct that was replication defective due to a deactivating mutation in the NS5B RNA polymerase. Already 24 hours after transfection a noticeable difference is evident in luciferase activity between the defective RNA and the 9-13F 10 or 5.1 constructs, while hardly any difference could be seen between the defective RNA (318 DN) and the parental RNA construct (wt) that possessed no adaptive mutations. During the whole period of observation, the highest 15 luciferase activity, and thus highest replication, was obtained with the 5.1-RNA. These results not only confirm the high replication efficiency of this RNA, but also show that it is possible to create a cell culture system with adapted HCV-RNA constructs for which the presence of a selectable gene is no longer necessary. A summary of the 20 nucleotide and amino acid differences between the starting construct and the mutants 9-13F, 5.1 and 19 is presented in Table 3.

25 **Example 8: Production of cell culture adapted HCV-RNA full-length genome**

In the examples 1 to 7 a subgenomic HCV-RNA was used which lacked the whole structural protein region from core up to 30 p7 or even NS2. It will be shown in this example that it is possible to make a HCV full-length genome replicate in cell culture with the aid of an adapted NS3-5B sequence. For this purpose the *Sfi*I fragment of the highly adapted HCV-RNA 5.1 produced according to Example 7 is first 35 transferred into a selectable HCV full-length genome (Fig. 12). This HCV genome was transfected into naive Huh-7 cells and subjected to selection with various G418

concentrations. Depending on the strength of selection (the G418 concentration), a varying large number of cell clones was obtained (Fig. 12 B). By contrast no colonies were obtained with the parental HCV full-length genome containing no adaptive mutations, as was the case for the negative control, which was replication defective due to a deactivating mutation in the NS5B RNA polymerase. To confirm that the thus resulting cell clones really contained an autonomously replicating HCV full-length construct, total RNA from several cell clones was isolated and analysed by means of the Northern Blot method. The full-length HCV-RNA was clearly detectable in all cell clones (Fig. 12). It is thus clearly confirmed, that with the aid of cell culture adapted HCV sequences it is possible to produce a HCV full-length genome, which replicates highly efficiently and autonomously in a cell line, i.e. adapted HCV full-length genomes can also be produced with the system of the invention. Furthermore, as this clone possesses the complete HCV sequence, i.e. it also possesses the structural proteins necessary for virus particle formation, it is possible to produce large quantities of infectious virus particles in cell cultures with this system. As a confirmation of these viruses, cell-free cell supernatants carrying a replicating HCV full-length genome, are added to naive Huh-7 cells and the thus infected cells subjected to selection with G418. Each cell clone growing under these conditions originates from an infected cell. The viruses in the cell culture supernatant of cells possessing a replicating HCV full-length genome can be enriched and purified using various known state of the art methods such as ultracentrifugation or microdialysis. They can then be used for the infection of naive cells. Using this method it is clearly demonstrated that cell culture adapted full-length genomes can be produced with the HCV cell culture system of the invention. These genomes replicate with high efficiency in cells and produce infectious viruses. The latter can be detected by

infection of an experimental animal, preferably a chimpanzee.

5    **Example 9: Production of HCV full-length constructs and HCV subgenomic constructs with reporter genes**

A HCV-RNA construct is produced in which a reporter gene is inserted in place of the antibiotic resistance gene (Fig. 13). Replication can thereby be determined through the quantity or activity of the reporter gene or reporter gene product. The reporter gene is preferably a gene from the group of the luciferase genes, the CAT gene (chloramphenicol acetyl transferase gene), the lacZ gene (beta galactosidase gene), the GFP gene (green fluorescence protein gene), the GUS gene (glucuronidase gene) or the SEAP gene (secreted alkaline phosphatase gene). This reporter gene and its products, namely the relevant reporter proteins, can be detected for example using fluorescence, chemiluminescence, colorimetrically or by means of immunological methods (for example, enzyme-linked immunosorbent assay, ELISA).

The reporter gene can be expressed either from a separate IRES or in the form of a fusion protein, which is active either as such or fused with a HCV protein via a proteolytically cleavable amino acid sequence in such a way that the reporter is separated from the HCV protein by cleavage of a cellular or viral (HCV) protease.

30

**Example 10: Production of HCV full-length constructs with integrated foreign genes used as liver cell specific vectors for gene therapy or as expression vectors.**

35

The construct (Fig. 14) is transfected in cells and leads to the formation of HCV virus particles that can be used

for the infection of further cells. Since the virus particles have encapsidated RNA with a foreign gene, it can be used in the infected cells for the production of the protein coded by this foreign gene. Cells transfected with  
5 the construct also express the foreign gene.

10 Example 11: Production of monocistronic HCV-RNA constructs  
in which the resistance gene product is  
expressed as a fusion protein with the HCV  
portion.

15 It is an advantage for some tests if the HCV-RNA construct does not possess a heterologous IRES element. Tests of this type are, for example, the determination of interferon  
resistance. If a cell possessing a HCV-RNA construct is  
incubated with interferon alpha or beta, a reduction in  
replication of the HCV-RNA results. In order to explain the  
mechanism of this effect it is necessary for the HCV-RNA  
20 construct not to possess any heterologous IRES, as  
otherwise it is not possible to determine whether the  
interferon mediated inhibition is via inhibition of the HCV  
replication or inhibition of the heterologous IRES. For  
this reason constructs are produced for which the  
25 resistance gene is fused with a HCV protein (Fig. 15).  
Either the fusion protein is active as such or the  
resistance gene product is linked to a HCV protein via a  
proteolytically cleavable amino acid sequence in such a way  
that it is separated from the HCV protein by a cellular or  
30 viral (HCV) protease.

**Table 1:** Specific infectivities (cfu/ $\mu$ g RNA) of HCV RNA constructs with adaptive mutations found with the 9-13F mutant introduced into the parental construct I<sub>389</sub>/NS3-3'/wt

amino acid exchange <sup>1</sup>	HCV protein	cfu/ $\mu$ g RNA <sup>2</sup>
none		30 - 60
1283 arg -> gly	NS3	200 - 250
1383 glu -> ala	NS3	30 - 60
1577 lys -> arg	NS3	30 - 60
1609 lys -> glu	NS3	160 - 300
(1283 arg -> gly + 1383 glu -> ala + 1577 lys -> arg + 1609 lys -> glu)	NS3	360 - 420
1936 pro -> ser	NS4B	1000-5000
2163 glu -> gly	NS5A	1000-5000
2330 lys -> glu	NS5A	30 - 60
2442 ile -> val	NS5A	30 - 60
all together		5000

5

<sup>1</sup> amino acid change in the polyprotein of the HCV isolate con 1 (EMBL-gene bank No. AJ238799); amino acids are given in single letter code.

<sup>2</sup> Colony forming units (number of cell clones) obtained with a selection of 500  $\mu$ g/ml G418.

**Table 2:** Specific infectivities (cfu/ $\mu$ g RNA) of the parental HCV RNA construct I<sub>389</sub>/NS3-3'/wt and the variants 9-13I, 9-13F, 5.1 and 19.

5

Transfected RNA variant	cfu/ $\mu$ g RNA <sup>1</sup>
wild type	30 - 50
9-13 I	100 - 1.000
9-13 F	1.000 - 10.000
5.1	50.000 - 100.000
19	50.000 - 100.000

<sup>1</sup> Colony forming units (number of cell clones) obtained with a selection of 500 $\mu$ g/ml G418.

**Table 3:** Nucleotide and amino acid sequence differences between the parental HCV RNA construct I<sub>389</sub>/NS3-3'/wt and the mutants 9-13I, 9-13F, 5.1 und 19

Clone	nt-position	nt-exchange	aa-exchange
9-13 I	3685	C > T	P > L
	4933	C > T	T > M
	5249	T > C	-
	8486	C > T	-
	8821	G > A	W > stop
	8991	C > G	R > G
	9203	A > G	-
	9313	T > C	F > S
9-13 F	9346	T > C	V > A
	3866	C > T	-
	4188	A > G	R > G
	4489	A > C	E > A
	4562	G > A	-
	4983	T > C	-
	5071	A > G	K > R
	5166	A > G	K > E
	6147	C > T	P > S
	6829	A > G	E > G
	7329	A > G	K > E
	7664	A > G	I > V
	8486	C > T	-
NK5.1	8991	C > G	R > G
	4180	C > T	T > I
	4679	C > T	-
	4682	T > C	-
	5610	C > A	L > I
	6437	A > G	-
	6666	A > G	N > D

	6842	C > T	-
	6926	C > T	-
	6930	T > C	S > P
	7320	C > T	P > S
	7389	A > G	K > E
NK19	3946	A > G	E > G
	4078	C > G	A > G
	4180	C > T	T > I
	4682	T > C	-
	5610	C > A	L > I
	5958	A > T	M > L
	6170	T > A	-
	6596	G > A	-
	6598	C > G	A > G
	6833	C > T	-
	6842	C > T	-
	6930	T > C	S > P
	7141	A > G	E > G
	7320	C > T	P > S
	7389	A > G	K > E
	7735	G > A	S > N

Given are the differences between the nucleotide and amino acid sequences of the parental HCV RNA sequence con 1 (EMBL-gene bank No. AJ238799) and those of the cell culture adapted HCV RNAs. Numbers refer to the nucleotide and amino acid positions of the con 1 isolate. nt, nucleotide; aa, amino acid.

## SEQUENCE LISTING

&lt;110&gt; Bartenschlager, Ralf

&lt;120&gt; Hepatitis C Virus cell culture system

&lt;130&gt; ba-1

&lt;140&gt; 199 15 178.4

&lt;141&gt; 1999-04-03

&lt;160&gt; 11

&lt;170&gt; PatentIn Ver. 2.1

&lt;210&gt; 1

&lt;211&gt; 11076

&lt;212&gt; DNA

&lt;213&gt; Hepatitis C Virus

&lt;400&gt; 1

gccagcccc gattgggggc gacactccac catagatcac tccccctgtga ggaactactg 60  
 tcttcacgca gaaagcgtct agccatggcg tragtatgag tgcgtgcag cctccaggac 120  
 cccccctccc gggagagcca tagtggctcg cggAACCGT gagtagacaccg gaattgccag 180  
 gacgaccggg tcctttcttg gatcaacccg ctcaatgcct ggagatttg gcgtgcccc 240  
 gcgagactgc tagccgaga gtgttgggtc gogaaaggcc ttgtggtaact gcctgatagg 300  
 gtgttgcga gtgcgggggg aggtctcgta gaccgtgcac catgagacag aatcctaaac 360  
 ctcaaagaaa aaccaaacgt aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420  
 acgcagggtc tccggccgct tgggtggaga ggctattcggt ctatgactgg gcacaacaga 480  
 caatcggtc ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttttt 540  
 ttgtcaagac cgacctgtcc ggtgcctgaa atgaactgca ggacgaggca gcgcggctat 600  
 cgtggctggc cacgacgggc gttccttgcg cagctgtgct cgacgttgacta actgaagcgg 660  
 gaagggactg gctgttattg ggcgaagtgc cggggcagga tctcctgtca tctcacctt 720  
 ctccctggca gaaagtatcc atcatgctg atgcaatgcg gcggctgtcat acgcttgatc 780  
 cggctacccg cccatttcgac caccaagcga aacatcgcat cgagcgcgca cgtactcgga 840  
 tggaaacccggt tcttgcgtat caggatgatc tggacgaaaga gcatcagggg ctgcgcggcag 900  
 cggactgtt cgccagggtc aaggcgcgc a tgcccgacgg caggatctc gtgtgaccc 960  
 atggcgatgc ctgttgcgc aatatcatgg tggaaaatgg ccgttttctt ggattcatcg 1020  
 actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttgct acccgtata 1080  
 ttgtctgaaga gcttggccgc gaatgggtc accgtttctt cgtgttttac ggtatcgccg 1140  
 ctcccgtattc gcagcgcattc gccttctatc gccttcttgc ctaggttccccc ccccccctt 1200  
 cagaccacaa cgggttccctt cttagcggat caattccgc cctctccctt ccccccctt 1260  
 aacgttactg gccgaagccg cttggaaataa ggccgggtgtg cttttgtata tatgttattt 1320  
 tccaccatatt tgccgttgc tggcaatgtg agggccggaa aacctggccc tgccttctt 1380  
 acgaggatcc cttaggggtct ttcccctctc gccaaaggaa tgcaggctt gttgaatgtc 1440  
 gtgaagaaag cagttcttctt ggaagcttctt tgaagacaaa caacgtctgt akgaccctt 1500  
 tgcaggcagc ggaacccccc acctggcgac'agggtcctt cggccaaaaa gccacgtgt 1560  
 taagatacac ctgcaaaaggc ggcacaaccc cagtgccacg ttgtgagttt gatagttgt 1620  
 gaaagagtca aatggcttc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680  
 gtacccatt gtatgggtc tgatctgggg cctcggtgca catgctttac atgtgttttag 1740  
 tcgagggttaaaaacgtctca gggcccccga accacggggg cttgttttcc tttgtaaaaa 1800  
 cacgataataa ccatgggcac gaaatctaaa cctccaaagaa aacacaaacg taacaccaac 1860  
 cgcccccac aggacgtcaa gttccgggc gttgggtcaga tcgtcggtgg agttacctg 1920  
 ttgcgcgcga gggcccccag gttgggtgtg cgccgcacta ggaagacttc cgagcggcgt 1980  
 caacctctgtt gaaaggcgaca acctatcccc aaggctcgcc agcccgagg tagggctgg 2040  
 gctcagcccg ggtacccctt gccccttat ggcaatgagg gttgggtgtt ggcaggatgg 2100  
 ctccctgtcac cccgtggctc tcggcctagt tggggcccca cggaccccg gcgttaggtcg 2160  
 cgcaatttgg gtaagggtcat cgataccctc acgtcggtct tcggcgatct catgggtac 2220  
 attccgtctcg tcggccggccccc cctagggggc gttggccaggcc ccttggcgca tggcgccgg 2280  
 gttctggagg acggcggtaa ctatgcacca gggaaatctgc cgggttgc ttttctatc 2340  
 ttccctttgg ctttgcgtc ctgtttgacc atccctagttt cccgttatga agtgcgcaac 2400  
 gtatccggag ttttgcgtt ctttgcgtt gggcggtgtt cccgttatgtt gacgggtc 2460  
 gcccgcacatgt ttttgcgtt ctttgcgtt gggcggtgtt cccgttatgtt gacgggtc 2520  
 cgctgctggg tagcgctcac tccacgctc gcccgcacatgtt cccgttatgtt gacgggtc 2580  
 acgataacac gccatgtca ttttgcgtt gggcggtgtt cccgttatgtt gacgggtc 2640  
 gtgggagatc ttttgcgtt ctttgcgtt gggcggtgtt cccgttatgtt gacgggtc 2700  
 cggcgcacatgtt ctttgcgtt gggcggtgtt cccgttatgtt gacgggtc 2760

cgtatggctt gggatatgat gatgaactgg tcacctacag cagccctagt ggtatcgca 2820  
 ttactccgga tcccacaagg tgcgtggat atggggcgg gggcccatgg gggagtcccta 2880  
 gcgggccttg cctactattc catgggggg aactgggcta aggttctgat tggatgtcta 2940  
 ctctttgcgg gctgtacgg gggacacct gracagggg ggacgtggc caaaaacacc 3000  
 ctcgggatc cgtccccctt ttcaccggg tcatccccaa aaatccaggt tgtaaacacc 3060  
 aacggcagct ggacatcaa caggactgc ctgaaactgca atgactccct caacactggg 3120  
 ttccttgctg cgctgttcta cgtgcacaag ttcaactcat ctggatgccc agagcgcatg 3180  
 gccagctgca gccccatcga cgcgttcgtc caggggtggg gggccatcac ttacaatgag 3240  
 tcacacagct cggaccagag gccttattgt tgcactacg caccgggccc gtgcgttata 3300  
 gtaccggcg cgcagggtgt tggccactg tactgcttca ccccaagccc tgcgtgggt 3360  
 gggacgaccg accgggtcg ggtccctacg tacagttggg gggagaatga gacggacgtg 3420  
 cgtcttcta acaaaccgcg ggcggccaa ggcacactgtt tggctgtat attgatgaat 3480  
 agcaactggt tcaccaagac gtggggggc cccccgtgtaa acatcggggg gatcggtcaat 3540  
 aaaaccttga cctgccccac ggactgttc cggaaagcacc cggaggccac ttacaccaag 3600  
 tgtggttcg ggccttgggt gacaccaga tgcttggtcc actaccata caggctttgg 3660  
 cactaccct gcaactgtcaa ctttaccatc ttcaagggtt ggtgtacgt ggggggagtg 3720  
 gagcacaggc tcgaaggccg atgcaattgg actcgaggag agcgttgtaa cttggaggac 3780  
 agggacagat cagagcttag cccgctgtc ctgtctacaa cggagttggca ggtattgccc 3840  
 tggcccttca ccaccctacc ggctctgtcc actgggttga tccatctcca tcagaacgtc 3900  
 gtggacgtac aataccctgtg cggataggg tggccgttgc aatcaaattgg 3960  
 gagtagtgtc tggctgttctt ctttcttcg gggacgcgc ggtctgtc ctgcttgg 4020  
 atgatgtgc tgatagctca agctgaggcc gccttagaga acctgggttgc cctcaacgcg 4080  
 gcatccgtgg cccggggcgc tggcatctc tccttcctcg tggcttctcg tgctgcctgg 4140  
 tacatcaagg gcaggctggt ccctggggcg gcatatgccc tctacggcgt atggccgcta 4200  
 ctccctgtcc tggccgtt accaccacga gcatacgcca tggaccggg gatggcagca 4260  
 tcgtgcggag ggcgggtttt cgtaggtctg atactcttga ctttgcacc gcactataag 4320  
 ctgttcttcg ctaggtctcat atgggttta caatatttt tcaccaggcc cgaggcacac 4380  
 ttgcaagtgt ggatcccccc cctcaacgtt cggggggggc gcatggccgt catcttcctc 4440  
 acgtgcgcga tccacccaga gtaatctt accatcacca aatctgtc cgccataactc 4500  
 ggtccactca tggcttcca ggctggata accaaagtgc cgtacttcgt gcgcgcacac 4560  
 gggctcatc gtgcattgcg gctggcggy aaggttctg ggggtcatta tgccaaatgg 4620  
 gctctcatga agtggccgc actgacaggt acgtacgtt atgaccatct caccctactg 4680  
 cgggactggg cccacgcggg cctacgagac ttgcgggtgg cagttgagcc cgctgtttc 4740  
 tctgatatgg agaccaaggat tatcacctgg gggcagaca cggccggcgtg tggggacatc 4800  
 atcttggggc tggccgtctc cggccgcagg gggaggggaga tacatctggg accggcagac 4860  
 agccttggaaag ggcagggggtg ggcactctc ggccttata cggctactc ccaacagacg 4920  
 cggggcttac ttggctgtcat ctttgcggc agctcgacc ttacttggt cacgaggcat 5220  
 gggggagggtcc aagtggtctc caccgcaaca caatcttcc tggccacccgtg cgtaatggc 5040  
 gtgtgttggc ctgtcttatca tggtgccggc tcaaaagaccc ttggccggcc aaagggccca 5100  
 atcaccctaaa tgcacaccaa tgggaccatg gacctcgatc gctggcaagc gccccccggg 5160  
 ggcgcgttctc tgacaccatg caccgtccgac agctcgacc ttacttggt cacgaggcat 5220  
 gccgatgtca ttccgggtcgcc cccgtctct actgtaaagggtt gggccacttctc gggccacgt 5340  
 gttggccatct ttccgggtcgcc cccgtcgatg ctatggaaac ctttgcggcgtt ggggggggtt 5400  
 cccggccgtac cgacagacatt ccagggtggcc catctacacg cccctactgg tagcggcaag 5520  
 agcaactaagg tggccgtctc gtatgcagcc caagggtata aggtgtttgt cctgaacccg 5580  
 tccgtcgccg ccacccctagg tttccggcg tataatgtcta aggcacatgg tattgcaccct 5640  
 aacatcgaaa ccggggtaag gaccatcacc acgggtgccc ccatcactgta ctccacctat 5700  
 ggcaagtttc ttggcgacgg tgggtgtctt gggggcgctt atgacatcat aatatgtat 5760  
 gagtgccact caactgactc gaccactat ctggccatcg gcacacttccctt gggccatctc 5820  
 gagacggctg gaggcgact ctttgcgtctc gccacccgtt cgcctccggg atccgttacc 5880  
 gtggccacatc caaacatcgaa ggagggtggct ctgtcccgatc ctggagaaat ccctttttat 5940  
 ggcaaaagcca tccccatcgaa gaccatcaag gggggggaggc accttcattt ctggccattcc 6000  
 aagaagaaat gtgtatgagct cggccgcggg ctgtccggcc tggactcaa tgctgttagca 6060  
 tattaccggg gccttgcgtt atccgtcata ccaactagcg gagacgtcat tgctgttagca 6120  
 acggacgctc taatgacggg ctttaccggc gatttcgact cagtgtatcgatc tgcacatata 6180  
 tgggtcaccct agacagtcga cttcagccctg gaccggaccc tcaccattga gaccacgacc 6240  
 gtggccacaaag acggcggtgc acgtcgatc cggccggggca ggacttggtag gggcaggatg 6300  
 ggcattttaca ggtttgtac tccaggagaa cggccctcg gcatgttgc ttctcggtt 6360  
 ctgtgcgtatg cttatgtacgc gggctgtct tggatcgacc tcacggccgc cgagacacta 6420  
 gttaggtgc gggcttaccc aaacacacca gggttggcccg tctgcccggg ccatctggag 6480  
 ttctgggaga gctgttttac aggccctacc caccatagacg cccatcttcc gtcccagact 6540  
 aaggcaggcag gagacaactt cccctacctg gtatgcatacc aggctacggt gtgcggccagg 6600  
 gctcaggctc caccctccatc gtgggaccaa atgtggaaat gtctcatacg gctaaagct 6660  
 acgctgcacg ggcacacgcg cctgctgtat aggctggag ccgttcaaaa cgaggtaact 6720  
 accacacacc ccataaccaa atacatcatg gcatgcattt cggctgaccc tggaggtcg 6780

acgagcacct gggtgctggt aggcggagtc cttagcagctc tggccgcgt a 6840  
 acaggcagcg tggtcattgt gggcagagtc atcttgcgtc gaaagccggc catcattccc 6900  
 gacagggaaag tcctttaccc ggagtgcgtg gagatggaaag agtgcgcctc acacccccc 6960  
 tacatcgaaac agggaaatgca gctcgccgaa caattcaaaac agaaggcaat cgggttgcgt 7020  
 caaacagcca ccaagcaagg ggaggctgtc gctcccggtt ggaatccaa gtggccgacc 7080  
 ctgcgaacct tctggccgaa gcatatgtt aatttcatac gcccggataca atattttagca 7140  
 ggcttgccca ctctgcctgg caaccccgat atagcatcac tgatggcatt cacagcctct 7200  
 atcaccagcc cgctcaccac ccaacatacc ctccctgtt acatccctggg gggatgggtg 7260  
 gcccggcaac ttgcgtccctc cagcgctgtc tctgcgttcc taggcggccgg catcgctgga 7320  
 gcccgtgttgc gcaagcatagg ccttggaaag gtgttgcgtt gatattttggc aggttatgg 7380  
 gcagggggtgg caggcgcgtc cgtggcctt aaggtcatga gcccggagat gcccctccacc 7440  
 gaggacttgg ttaacactt ccctgttact ctctccctgt gccccttgcgt cgtcgggggtc 7500  
 gtgtgcgcag cgatactgcg tcggcactgt gcccgggggg aggggggtgt gcagtggatg 7560  
 aaccgggtga tagcgttgc ttcgggggtt aaccacgtct ccccccacgc ctagtgcct 7620  
 gagagcgacg ctgcagcagc tgtcactcag atccctctcta gtcttaccat cactcagctg 7680  
 ctgaagaggc ttacccatgt gatcaacgg gactgttccca cccatgtc cggctcgagg 7740  
 ctaagagatg tttgggattt gatatgcacg gtgttgcgtt atttcaagac ctggctccag 7800  
 tccaagctcc tggccgcatt gcccggatc cccttcttct catgtcaacg tgggtacaag 7860  
 ggagtctggc gggggcgcacgg catcatgca accacctgca catgtggagc acagatcacc 7920  
 ggacatgtga aaacacggttc catgaggatc gtggggctt ggacctgttag taacacgtgg 7980  
 catggacat tccccctaa cgcgtatacc acggggccctt gcaacgcctc cccggcgcac 8040  
 aattatcttca gggcgctgtt gcccgggtt gctgaggagt acgtggaggat tacgcgggtg 8100  
 ggggatttcc actacgttgc gggcatgacc actgacaacg taaagtgcctt gtgtcagggtt 8160  
 cccggccccc aatttttca gagaagtggat ggggtgcgtt tgcacaggtt cgcctcagcg 8220  
 tgcaaaccctt ccctacggg gggaggttaca ttccctgggtc ggctcaatca atacctgggtt 8280  
 gggtcacagc tcccatgcg gcccgaaccg gacgttagcag tgctcacttc catgtctacc 8340  
 gaccctccca acattacggc gggagacggct aagcgtaggc tggccagggg atctccccc 8400  
 tccttgcaca gtcatcagc tagccagctg tctgcgttccctt ccttgaaggc aacatgcact 8460  
 accccgtatc actcccccggg cgtgcgttcc atcgagggttcc accttctgtt gggcaggag 8520  
 atggggcggtt acatcaccgg cgtgggttca gaaaataagg tagtaatttt ggactctttc 8580  
 gagccgttcc aaggcgaggaa ggatggaggaa gaatgttccg ttccggcgga gatccctgg 8640  
 aggtccaggaa aattccctcg agcgatgttttccatggcactt gcccggatttta caaccctcca 8700  
 ctgttagagt ccttggaaaggc cccggactac gtccctccag tggtacacgg gtgtccattt 8760  
 cccgcgttcca aggccccctcc gataccaccc tccacggagga agaggacggt tgcgtgtca 8820  
 gaatctaccg tgcgttctgc ctggccggag ctcgcaccaa agaccttcgg cagctccgaa 8880  
 tcgtcgcccg tcgacagcg gacggccacg gctcttcgtt accagccctc cgacgacggc 8940  
 gacgcgggat ccgacgttgc gtcgtactcc tccatggccc cccttgggggg ggagccgggg 9000  
 gatccccccatc tcaacggcgg gtcgttgcgttccatggcactt acgtaatggcaggatggtagt tgaggacgtc 9060  
 gtctgtgtct cgtgttccatc cacatggaca ggcgccttgc tcaacggccatc cgctcgccgg 9120  
 gaaaccaagc tgcccatcaaa tgcactgagc aacttcttgc tccgttccatca caacttgggtc 9180  
 tatgttacaa catctcgatc cgcaaggctt cggcagaaga aggttacccctt tgacagactg 9240  
 caggcttccgg acgaccacta ccgggacgttca ctaaggaga tgaaggcgaa ggcgtccaca 9300  
 gttaaaggcta aacttcttccatc cgtggaggaa gctgttacgc tgacgcccccc acattggcc 9360  
 agatctaaat ttggctatgg ggcacaaaggac gtcggaaacc tatccagcaaa ggcgttaac 9420  
 cacatccgttcc ccttgcgttcc gggactgttgc gaaagactgtt agacaccaat tgacaccacc 9480  
 atcatggcaaa aaaatgggtt tttctgttccatc caaccagaga agggggccgg caagccagct 9540  
 cgccttccatc tattccca gttttgggtt cgtgtgttgc gggaaaatggc cctttacat 9600  
 gtggcttcca ccctccctca ggcgtgtat ggccttccat acggatttca atactcttcc 9660  
 gagacagccgg tcgagttcc ggtgaatgcc tggaaagcgaa agaaatggcc tatgggttcc 9720  
 gcatatgaca cccgcgtttt tgacttcaacg gtcactgaga atgacatccg tgggttgggg 9780  
 tcaatctacc aatgttgc tttggccccc gaaaggccagac agggccataag gtcgttccatca 9840  
 gagcggctttt acatcgggggg ccccttactt aattctaaag ggcagaactt cggctatcgc 9900  
 cgggtcccgccg cgaggcggtt actgtacggacc agtgcggatc atacccttccat atgttactt 9960  
 aaggcccgctg cggccgttcc agtgcgttcc ctcggactt gcaacgtgttgcgttccatc 10020  
 gacgacccgttcc tgcgttccatc tggaaaggcg gggacccaaaggacggaggc gacccatccgg 10080  
 gccttccatcc aggttccatc tagatactt gcccggccctt gggacccggcc caaaccagaa 10140  
 tacgacttgg agttccatc atcatgttcc tccaaatgttgc tgcgttccatc cgtgcgttccatc 10200  
 ggccaaaagggtt tgcgttccatc caccgggttccatc cccaccaccc cccttgcgttccatc ggtgttgcgttcc 10260  
 gagacagacta gacacactcc agtcaatcc tgggttccatc acatcatcat gatgcgccttcc 10320  
 accttgcgtttt caaggatgtt cctgttccatc catttcttccatc ccaccccttccatc agtgcaggaa 10380  
 caacttgcgtttt aaggcccttccatc ttgttccatc tggggccctt gttacttccat tggccactt 10440  
 gacccatccatc agatccatc acgacttccatc ggccttccatc catttcttccatc ccatgttccatc 10500  
 tctccagggttccatc ggtgttccatc tgcgttccatc aacttgggggtt accggcccttccatc 10560  
 cgagtctggc gacatcgccggc cagaatgttccatc cggccgttccatc tactgttccatc ggggggggg 10620  
 gtcgttccatc gtggcaagttt ccttccatc tgggttccatc tgcgttccatc ttacagccgg 10740  
 ccaatcccggtt ccgttccatc gttggatccatc tccagttccatc tgcgttccatc ttacagccgg 10800  
 ggagacatcatc atcacagccatc gtctcgccatc cgaccccttccatc gtttccatc gtcgttccatc 10800

CA 02303526 2000-03-31

ctactttctg taggggttagg catctatcta ctccccaacc gatgaacggg gagctaaaca 10860  
ctccaggcca ataggccatc ctgtttttt ccctttttt ttttctttt tttttttttt 10920  
ttttttttt ttttttttc tcctttttt ttccctttt ttccctttc ttccctttgg 10980  
tggctccatc ttagccctag tcacggctag ctgtgaaagg tccgtgagcc gcttgactgc 11040  
agagagtgtct gatactggcc tctctgcaga tcaagt 11076

<210> 2  
<211> 8637  
<212> DNA  
<213> Hepatitis C Virus

<400> 2

gccagcccc	gattggggc	gacactccac	catacatcac	tcccctgtga	gaaactactg	60
tcttcacgca	gaaagcgtct	agccatggcg	ttatgtatgag	tgtcgtgcag	cctccaggac	120
ccccccccc	gggagagcca	tagtggctcg	cgaaaccgg	gagtacaccc	gaattgccag	180
gacgaccgg	tccittctt	gatcaaccc	ctcaatgcct	ggagatttg	gcgtgcccc	240
gcgagactgc	tagccgagta	gtgttgggtc	gcggaaaggcc	ttgtggta	gcctgatagg	300
gtgttgcga	gtccccccgg	aggctctcg	gaccgtgcac	catgagcacg	aatccaaac	360
ctcaaaaagaa	zaccaaagg	cgcgcattga	ttgaacaaga	tggatgcac	cgaggttctc	420
cggccgc	gttgtggaggg	ctattcgct	atgtatggc	acaacagaca	atcggtctgt	480
ctgatgcgc	cgtgtccgg	ctgtcagcgc	agggcgccc	ggttctttt	gtcaagacccg	540
acctgtccgg	tgcctgaat	gaactgcagg	acgaggcagc	gcggctatcg	tggctggca	600
cgacggcg	tccttgcga	gctgtgc	acgttgtac	tgaagcggga	aggactggc	660
tgctattgg	cgaagtgc	gggcaggatc	tcctgtc	tcaccttgct	cctgcccaga	720
aagtatccat	catggctgtat	gcaatgcgc	ggctgcatac	gcttgatccg	gctacactgc	780
cattcgacca	ccaaagc	catcgcatcg	agcagagcacg	tactcgatg	gaagccggc	840
ttgtcgatca	ggatgatctg	gacaaagagc	atcaagggg	cjcgccagcc	gaactgttc	900
ccaggctcaa	ggcgcgc	cccgcacgg	aggatctcg	cgtgaccat	ggcgtgcct	960
gcttgcggaa	tatcatgg	aaaaatggc	getttctgg	attcatcgac	tgtggccggc	1020
tgggtgtgg	ggaccgtat	caggacatag	cggtggctac	ccgtgatatt	gctgaagagc	1080
ttggcggcg	atgggtgc	cgcttc	tgcttacgg	tatcgccgt	cccgattcgc	1140
agcgcatgc	cttctatgc	cttcttgac	agttttctg	agtttaaaca	gaccacaacg	1200
gtttccctct	agcgggatca	atccgc	tctccctcc	ccccccctaa	cgttactggc	1260
cgaagccgt	tggataaagg	ccgggtgtcg	tttgttata	tgttatttc	caccatattg	1320
ccgtctttt	gcaatgtgag	ggcccgaaa	cctggccc	tcttcttgac	gagattcct	1380
agggtctt	ccccctctcg	caaaggatg	caaggtctgt	tgaatgtcg	gaaggaagca	1440
gttctctgg	aagcttctg	aagacaaca	acgtctgtag	cgacccttg	caggcagcgg	1500
aaccccccac	ctggcgcac	gtgcctctgc	ggccaaaagc	cacgtgtata	agatacacct	1560
gcaaaggcg	cacaacccca	gtgccac	gtgagttga	tagttgtga	aagagtcaaa	1620
tggctctct	caagcgatt	caacaagg	ctgaaggatg	cccagaaggt	acccattgt	1680
atgggatctg	atctggggcc	tcggtgca	tgcttacat	gtgttagtc	gaggtaaaa	1740
aacgtctagg	ccccccgaa	cacggggac	tggtttct	ttgaaaaca	cgataatacc	1800
atggaccgg	agatggcage	atcg	ggcgggtt	tctgttagt	gataacttt	1860
accttgtc	cgcata	gtgttc	ctcaggctca	tatgtgtt	acaatattt	1920
atcaccaggg	ccgaggcaca	cttgca	tgatcccc	ccctcaacgt	tgggggggc	1980
cgcgtatgc	tcatcctct	cacgtgc	atccaccc	agctaattt	taccatcacc	2040
aaaatctgc	tcgcata	cggtccactc	atgtgtctc	aggctgtat	aaccaaagt	2100
ccgtactcg	tgcgc	cggtctcatt	cgtcatgc	tgcgtgtcg	gaaggttgc	2160
gggggtcatt	atgtccaaat	ggctctcat	aagg	cactgacagg	tacgtac	2220
tatgaccatc	tcacccact	gggggactgg	gcccacgcgg	gcctacgaga	ccttgcgg	2280
gcagttgagc	ccgtgtctt	ctctgtat	tgagacagg	ttatcac	ggggcagac	2340
accgcggcg	gtggggatc	catctgg	tttgcgtt	ccgcccgc	ggggagggag	2400
atacatctgg	gaccggcaga	cagcttga	ggcaggggt	ggcacttct	cgcgcatt	2460
acggctact	cccaacagac	gcgaggc	cttggctca	tcatcactag	cctcacaggc	2520
cgggacagga	accaggtcg	gggggaggtc	caagtgg	ccaccgc	acaatttt	2580
ctggcgcac	gctgtcaatgg	cgtgtgtt	actgtctatc	atggtgc	ctcaaagacc	2640
cttgcgc	caaaggccc	aatcacc	atgtacacca	atgtggacca	ggacactgc	2700
ggctggcaag	cgccccccgg	ggcgcgtt	ttgacacat	gcacactgc	cagctcg	2760
cttacttgg	tcacaggca	tgccgtat	atccgggtc	ggccgcgg	cgacagcagg	2820
ggggcctac	tctccccc	gcccgttcc	tacttgaagg	gtcttctgg	cgttccact	2880
ctctgc	cggggcac	tgtggc	tttgcgtt	cgtgtgc	ccgagggtt	2940
gcgaaggcg	tggactttgt	acccgtcg	tctatggaa	ccactatgc	gtccccgtc	3000
ttcacggaca	actcgcccc	tccggcgt	ccgcagacat	tccagg	ccatctacac	3060
gcccctactg	gtagcggca	gagacta	gtgcggct	cgtatgc	ccaagggtat	3120
aagggtctt	tcctgaaccc	gtccgtc	gcacccctag	gttgcgg	gtatatgtct	3180
aaggcacatg	gtatcgaccc	taacatc	accggggtaa	ggaccat	cacgggtgc	3240
cccatcacgt	actccacca	tggca	cttgcgc	gtgggtgc	tgggggc	3300
tatgacatca	taatatgt	tgatgtc	actgtact	cgaccat	ccttggc	3360
ggcacagtc	tggacca	ggagacg	ggagcgc	tcgtcg	cgccaccgc	3420
acgcctccgg	gatcggtac	cgtgc	ccaaacatc	aggagg	tctgtcc	3480
actggagaaa	tccctttt	tggcaagcc	atccccatc	agaccat	gggggggagg	3540
cacctcattt	tctgc	caagaagaaa	tgtgtatg	tcgcgcg	gctgtcc	3600

tcgcgactca	atgctgttagc	atattaccgg	ggcccttgatg	tatccgtcat	accaactagc	3660
ggagacgtca	ttgtcgtagc	aacggacgct	ctaatacgacgg	gctttaccgg	cgatttcgac	3720
tcaagtatcg	actgcaatac	atgtgtcacc	cagacagtcg	acttcagcct	gaccccgacc	3780
ttcaccattg	agacgacgac	cgtgccacaa	gacgcccgtgt	cacgctcgca	gcggcgagge	3840
aggactggta	ggggcaggat	gggcattac	aggtttgtga	ctccaggaga	acggccctcg	3900
ggcatgttcg	attcttcggt	tctgtcgtag	tgtatgacgg	cgggctgtc	ttggatcag	3960
ctcaccccc	ccgagaccc	agtttagttt	cgggcttacc	taaacacacc	agggttgccc	4020
gtctgccagg	accatctgg	gttctgggg	agcgttta	caggctcac	ccacatagac	4080
gcccattct	tgtcccagac	taagcaggca	ggagacaact	tcccctacct	gttagcatac	4140
caggctacgg	tgtgcgccag	ggctcaggct	ccacccat	cgtgggacca	aatgtggaaag	4200
tgtctcatac	ggctaaagcc	tacgctgcac	ggccaacgc	ccctgtgt	taggctggaa	4260
gcccgtcaaa	acgaggttac	taccacacac	ccataacca	aatacatcat	gcatgcat	4320
tcggctgacc	tggaggtcgt	cacgaccc	tgggtgtctgg	tagggaggt	cctagcaggt	4380
ctggccgcgt	attgcctgac	aacaggcage	gtggteattt	ttggcaggat	cattttgtcc	4440
ggaaaagccgg	ccatcatcc	cgacaggaa	gtctttacc	gggagttcga	ttagatggaa	4500
gagtgcgcct	cacacccccc	ttacatcgaa	caggaaatgc	agctcggca	acaattcaaa	4560
cagaaggcaa	tcgggttgc	gcaaacagcc	accaagcaag	cggaggctgc	tgctccctgt	4620
gttgaatcca	agtggcggac	cctcgaagcc	ttctggcga	agcatatgt	gaatttcatc	4680
agcgggatac	aatattttac	aggcttgc	actctgcctg	gcaacccgc	gatagcatac	4740
ctgatggcat	tcacagcctc	tatcaccagc	ccgctcacca	cccaacatac	cctctgttt	4800
aacatctgg	ggggatgggt	ggccgcacaa	cttgctcc	ccagcgtgc	ttctgtttc	4860
gttaggcgcct	gcatcgctgg	agcgggtt	ggcagcatag	gccttggaa	gtgttgcgt	4920
gatattttgg	caggttatgg	agcagggtt	ggaggcgcgc	tctgtgcctt	taaggtcata	4980
agcggcggaga	tggccctccac	cgaggac	gttaacactac	tccctgtat	cctctccct	5040
ggcgccttag	tcgtcggtt	cgtgtcgca	gcatactgc	gtcggcacgt	ggcccccagg	5100
gagggggctg	tgcaatggat	gaaccggctg	atagcgttc	ctrccgggg	taaccacgtc	5160
tccccccacgc	actatgtgc	tgagagcgac	gtgcagcac	gtgtcaact	gatcctctc	5220
agtcttacca	tcactcagct	gctgaagagg	cttcaccagt	ggatcaacga	ggactgtcc	5280
acgccatgtct	ccggctcg	gctaagagat	gtttggat	ggatgtc	gtgttgc	5340
gatttcaaga	cctggctcca	gttcaagctc	ctggccgcgt	tgccggaggt	cccccttetc	5400
tcatgtcaac	gtgggttacaa	gggagcttgg	cggggcacg	gcatatcgca	aaccacgtc	5460
ccatgtggag	cacagatcac	cgagacatgt	aaaaacgggt	ccatgaggat	ctgtgggcct	5520
aggacctgt	gttaacacgt	gcatggaaaca	tttccccatta	acgcgtacac	cacgggcccc	5580
tgcacccct	ccccggcgcc	aaattatttct	agggcgctgt	ggcgggtggc	tgctgaggag	5640
tacgtggagg	ttacgggtt	gggggattt	cactacgt	cgggcatgac	cactgacaac	5700
gtaaagtgc	cgtgtcaggt	tccggcccc	gaattcttca	cagaagtgg	ttgggtgcgg	5760
ttgcacaggt	acgctccagc	gtgcaaaaccc	cttctacggg	aggaggtc	acttctgtt	5820
gggctcaatc	aatacctgg	tgggtcacag	ctcccatgc	agcccaacc	ggacgttaga	5880
gtgctcaatt	ccatgtc	cgacccctc	cacattacgg	cggagacggc	taagcttgg	5940
ctggccagg	gatctcccc	ctcttggcc	agtcatcag	ctagccag	gtctgcgcct	6000
tccttgcagg	caacatgcac	tacccgtcat	gactccccgg	acgctgac	catcgaggcc	6060
aacctctgt	ggcgccagga	gatgggggg	aacatcacc	gcgtggagtc	agaaaataa	6120
gttagtaattt	tggactcttt	cgagccg	caagcggagg	aggatgagag	gaaagtatcc	6180
gttccggcgg	agatctcg	gaggtccagg	aaattccctc	gagcgtat	catatggca	6240
cgccccggatt	acaacccctcc	actgttag	tcctggaaagg	acccggacta	ctggccctca	6300
gtggtacacg	ggtgtccatt	ggccctg	aaggccctc	cgatacc	tccacggagg	6360
aagaggacgg	tttgcctgtc	agaatctacc	gtgttcttct	ccttgcgg	gtctgcgcaca	6420
aagacccct	gcaatcg	atcgtcg	gtcgaac	gcacggcaac	ggcctctc	6480
gaccggccct	ccgacgacgg	cgacgcgg	tccgacgtt	agtcgtact	ctccatgc	6540
cccccttgcagg	gggagccggg	ggatcccgat	ctcagcgac	gttctgg	taccgtaa	6600
gaggaggcta	gtgaggacgt	cgtctgtc	tcatgtt	acacatggac	aggccctcg	6660
atcacccat	gctgtcg	ggaaacca	ctgccc	atgcactg	caactctt	6720
ctccgtcacc	acaacttgg	ctatgtc	acatctcg	gcgcac	ggggcagaa	6780
aaggtcacct	ttgacagact	gcaggctt	gacgacc	acccggac	gtcaaggag	6840
atgaaggcga	aggcgtccac	agttaa	aaaaacttata	ccgtgagga	agcttgc	6900
ctgaccccc	cacattcg	cagat	tttggat	ggcgaagg	ctgtccggaa	6960
ctatccagca	aggccgtt	ccacatcc	tccgtgtt	aggacttgc	gaaagacact	7020
gagacacca	ttgacacc	catcatgg	aaaaatgagg	ttttctgc	ccaaccagag	7080
aagggggggc	gcaagcc	tcgcctt	gtattcc	atttgggg	tcgtgttgc	7140
gagaaaatgg	cccttacg	tgtggtct	accctcc	agggcgt	ggctcttca	7200
tacggattcc	aatactctcc	tggacag	gtcgagtt	tggat	ctggaaagcg	7260
aagaaatg	ctatggc	cgcatat	accgcgtt	ttgact	caactcg	7320
aatgacatcc	gtgttgg	gtcaat	caatgtt	acttggcc	ccaaaggcaga	7380
caggccat	gtgtcg	agagcg	tacatcg	gccccctg	taattctaa	7440
ggcgaact	ggcgtc	ggcgtc	ggcgtt	tactgac	cagctgcgt	7500
aataccctca	catgttact	gaaggcc	ggggctgt	gagctgc	gtccaggac	7560
tgcacqat	tcgtatgc	agacgac	gtcgttat	gtgaaagcg	ggggacccaa	7620

gaggacgagg cgagcctacg ggccttcacg gaggttatga ctagatactc tgccccccct 7680  
ggggaccgc ccaaaccaga atacgactt gaggatgaa catcatgtc ctccaatgtg 7740  
tcagtcgcgc acgatgcac tggaaaagg gtgtactatc tcacccgtga ccccaccacc 7800  
ccccttgcgc gggctgcgtg ggagacact agacacactc cagtcaattc ctggctaggc 7860  
aacatcatca tgtatgcgc caccttggtt gcaaggatga tcctgtatgc tcatttcttc 7920  
tccatccctc tagtcagga acaacttggaa aaagccctag attgtcagat ctacggggcc 7980  
tgttactcca ttgagccact tgacctaccc cagatcattc aacgacttca tggccttagc 8040  
gcattttcac tccatagttt ctctccaggat gagatcaata gggtggcttc atgcctcagg 8100  
aaacttgggg taccgcctt gcgagtctgg agacatcggtt ccagaagtgt ccgcgcctagg 8160  
ctactgtccc agggggggag ggctgccact tggcaagt acctcttcaa ctgggcagta 8220  
aggaccaagc tcaaactcac tccaatcccg gctgcgtccc agttggattt atccagctgg 8280  
ttcggttgcgtg gttacagcg gggagacata tatacagcc tgcgttgcgtgc ccgaccccg 8340  
tggttcatgtt ggtgcctact cctacttttctt gttagggtagt gcatctatct actccccaaac 8400  
cgatgaacgg ggagctaaac actccagggcc aataggccat cctgtttttt tccctttttt 8460  
tttttctttt tttttttttt tttttttttt tttttttttt ctccctttttt ttccctctttt 8520  
ttttccctttt ctccctttt gttggctccat cttagcccta gtcacggcta gctgtgaaag 8580  
gtccgtgagc cgcttgcactg cagagatgc tgatactggc ctctctgcag atcaagt 8637

<210> 3  
<211> 8001  
<212> DNA  
<213> Hepatitis C Virus

<400> 3

```

gcccggccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgcgtgcag cctccaggac 120
ccccctcccc gggagagcca tagtggctcg cgaaaccggc gaggaccccg gaattgccag 180
gacgaccggg tcctttctt gatcaacccg ctaatgcct ggagatttg gctgcccccc 240
gcgactgc tagccgatgt gtgtgggtc gcaaaggcc ttgtgtact gcctgatagg 300
gtgctgcga gtggccccc aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaaacgt aacaccaacg ggccgcctat gattgacaa gatggattgc 420
acgcagggttcc tccggccgtc tgggtggaga ggctattcggt ctatgactgg gcacaacaga 480
caatcggtcg ctctgatgcc gccgtgttcc ggctgtcagc gcagggcgc cccgttctt 540
ttgtcaagac cgacctgttcc ggtgcctgtaa atgaaactgca ggacgaggca gcgcggctat 600
cgtggctggc cacacggggc ttcccttgcg cagctgtgtc cgacgttgtc actgaagcgg 660
gaaggggactg getgtatgg ggcgaatgc cggggcagga ttcctgtc ttcacccgt 720
ctccgtccga gaaagtatcc atcatgtcg atgcaatgcg gcgctgtcat acgcttgatc 780
cggttacccgt cccatttcgac cacaaggcga aacatcgcat cgagcgagca cgtactcgga 840
tggaaagccgg tcttgcgtat caggatgtc tggacgaa gcatcagggg ctcgcggccag 900
ccgaactgtt cgccaggctc aaggcgcgcgca tgcccgcacgg cgaggatctc gtctgtgaccc 960
atggcgatgc ctgttgcgg aatatcatgg tggaaaatgg ccgttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcgaccgtc atcaggacat agcgttggtc acccggtata 1080
ttgtctgaaga gctggccgc gaatgggtcg accgttctt cgtgttttac ggtatcgccg 1140
ctcccgatcc gcaacgtatc gccttctatc gccttctgtaa cgatgttcc tgagttaaa 1200
cagaccacaa cgggttccctt atagcggat caatccgcgc ccttccctc cccccccct 1260
aacgttactg gccgaagccg tccaccatata tgcgttccat acgagcatc ctaggggtctt gttggatataa ggccgggtgtg cgtttgtcta tatgttattt 1320
tggcaatgtg aggccccggaa aacctggccc tgcgttccatc gcaagggtct gttgaatgtc 1380
tcccccttc gccaaaggaa tgcaagggtct tgaagacaaa caacgtctgt agcggaccctt 1440
gtgaaggaag cagttctctt ggaagcttctt tgatctgggg cctcgggtgcgcatgttccatc tttttttc ctttggaaaa 1500
tgcaggcagc ggaacccccc taagatacac ctgcaaaaggc gaaagatca aatggctctc gtggggatataa ggccgggtgtg cgtttgtcta tatgttattt 1560
tgcgttccatc tggcaatgtg aggccccggaa aacctggccc tgcgttccatc gcaagggtct gttgaatgtc 1620
gttccaccatc gtatgggttcaatc gcaacacaatc tttccctggcg acctgcgtca atggcgtgtg ttggactgtc 1680
tatcatgttgc cggctcaaa gacccttgcg ggcccaaagg gccaatcac ccaaatgtac 2040
accaatgtgg accaggaccc ccatgcacct gcccggccgc gggcggacagc aagggtctt cgggggttcc gctggcgtgt gcaacggggcggccgc ttccttgaca 2100
ggcacaaccc cagtgccacg ctcaagggttcaatc gcaacacaatc tttccctggcg acctgcgtca atggcgtgtg ttggactgtc 2160
gttccaccatc gtatgggttcaatc gcaacacaatc tttccctggcg acctgcgtca atggcgtgtg ttggactgtc 1920
gttccaccatc gtatgggttcaatc gcaacacaatc tttccctggcg acctgcgtca atggcgtgtg ttggactgtc 1980
tatcatgttgc cggctcaaa gacccttgcg ggcccaaagg gccaatcac ccaaatgtac 2040
cgtcggctgg caagcgcggcc ggcacccatc ttgtgtacgc gcaacacaatc tttccctggcg acctgcgtca atggcgtgtg ttggactgtc 2160
ggccggccgc gggcggacagc aagggtctt cgggggttcc gctggcgtgt gcaacggggcggccgc ttccttgaca 2220
gttccaccatc gtatgggttcaatc gcaacacaatc tttccctggcg acctgcgtca atggcgtgtg ttggactgtc 2280
gatgtatccatc gcaacacaatc tttccctggcg acctgcgtca atggcgtgtg ttggactgtc 2340
gaaaccacta tgcgttccatc ggttccatc gacaactctgt cccctccggc cgtacccgtcag 2400
acattccagg tggcccatct acacgccttactgttagcg gcaagagcac taaggtgccg 2460
gttccaccatc gtatgggttcaatc gcaacacaatc tttccctggcg acctgcgtca atggcgtgtg ttggactgtc 2520
ctagggttccatc gcaacacaatc tttccctggcg acctgcgtca atggcgtgtg ttggactgtc 2580
gttccaccatc gtatgggttcaatc gcaacacaatc tttccctggcg acctgcgtca atggcgtgtg ttggactgtc 2640
gttccaccatc gtatgggttcaatc gcaacacaatc tttccctggcg acctgcgtca atggcgtgtg ttggactgtc 2700
gttccaccatc gtatgggttcaatc gcaacacaatc tttccctggcg acctgcgtca atggcgtgtg ttggactgtc 2760
gttccaccatc gtatgggttcaatc gcaacacaatc tttccctggcg acctgcgtca atggcgtgtg ttggactgtc 2820
atcgaggggg tggctctgtc cggcttccatc acgtactccatc cctatggcaa gtttcttgcc 2880
atcgaggggg tggctctgtc cggcttccatc acgtactccatc cctatggcaa gtttcttgcc 2940
gagctcggcc gcaagctgtc cggcttccatc ctcaatgtcg tagcatatata ccggggccctt 3000
gatgtatccatc tcataccaaatc tagcggagac gtcttgcgtcg tagcaacggaa cgctctaattg 3060
acgggttccatc cggcgattt cggcttccatc acgtactccatc cctatggcaa gtttcttgcc 3120
gttccaccatc gcaacacaatc tttccctggcg acctgcgtca atggcgtgtg ttggactgtc 3180
gttccaccatc gcaacacaatc tttccctggcg acctgcgtca atggcgtgtg ttggactgtc 3240
gttccaccatc gcaacacaatc tttccctggcg acctgcgtca atggcgtgtg ttggactgtc 3300
gacggggcttcaatc gcaacacaatc tttccctggcg acctgcgtca atggcgtgtg ttggactgtc 3360
tacctaaaca caccagggtt gcccgttccatc cggaccatc tggagttctg ggagagcgtc 3420
tttacaggcc tcaccacat agacgcccatttccatc agactaagca ggcaggagac 3480
aacctccctt acctggatcg ataccaggatc acgggtgtcg ccagggttccatc ggctccaccct 3540

```

ccatacggtgg accaaaatgtg gaagtgtctc atacggctaa agcctacgct gcacgggcca 3600  
 acggccctgc tgrataggtgc gggagccgtt caaaacgagg ttactaccac acaccccata 3660  
 accaaataca tcattggcatg catgtcgct gacctggagg tcgtcacagag cacctgggtg 3720  
 ctggtaggcg gagtcctagc agctctggcc gcgtattgcc tgacaacagg cagcgtggc 3780  
 attgtggca ggatcatctt gtccggaaag ccggccatca tttccgcacag ggaagtccctt 3840  
 taccggggatg tgcgtatggat ggaagagtgcc gcctcacacc tcccttatcat cgaacaggga 3900  
 atgcagctcg ccgaaacatt caaacagaag gcaatcggtg tgctgcaaac accaccaag 3960  
 caaaggagg ctgtgtctcc cggtgtggaa tccaaatgtgc ggaccctcgaa gacctctgg 4020  
 gcgaagcata tggaaatatt catcagcccc atacaatatt tagcaggctt gtccactctg 4080  
 cctggcaacc ccgcgtatgc atcactgtat gcattcacag cctctatcac cagccccgtc 4140  
 accacccaaac ataccctctt gtttaacatc ctggggggat gggtggccgc ccaacttgct 4200  
 cctcccaagcg ctgtctgtgc ttctgttaggc gccggcatcg ctggagccgc tggggcagc 4260  
 ataggccttg ggaaggtgtc tggatatt ttgcgtggat atggagcagg ggtggcaggc 4320  
 ggcgtcggtt ccttaaggt catgagccgc gagatggccct caccggagga cttggtaac 4380  
 ctactccctg ctatcccttc ccttggccgc ctatgtcg gggtctgtg cgcacgtata 4440  
 ctgcgtccgc acgtggggcc aggggagggg gctgtcagt ggtatggatc gctgtatgc 4500  
 ttgcgttcgc gggtaacca cgtctccccc acgcactatg tgccctgagag cgacgtcgca 4560  
 gcacgtgtca ctcaagatctt ctctagttt accatcactc agctgtcaa gaggtttcac 4620  
 cagtggatca acgaggactg ctccacggca tgctccggct cgtggtaaag agatgtttgg 4680  
 gattggatat gcacgggtt gactgtatcc aagacctggc tccagtc当地 gctccctggc 4740  
 cgattgcgg gactccctt ctctctatgt caacgtgggtt acaaggagg ctggggggc 4800  
 gacggcatca tgeaaaccac ctggccatgt ggacacaga tgccggaca aatattcccc 4860  
 gtttccatga ggatcggtt gcttaggacc tggatgttgggg tggatggatc aacattcccc 4920  
 attaacgcgt acaccacggg cccctgcacg ccctccccgg cgccaaatta ttcttagggcg 4980  
 ctgtggccgg tggctgtga ggagtacgt gagggtacgc gggtggggta tttccactac 5040  
 gtgacggca tgaccactga caacgtaaag tgcccgtgtc aggttccggc ccccaatttc 5100  
 ttcacagaag tggatgggg tggcgtgcac aggtacgtc cagcgtgca acccttccta 5160  
 cgggaggagg tcacattctt ggtgggtc aatcaatacc tgggtgggtc acagtc当地 5220  
 tgcgagcccg aaccggacgt agcgtgtc actccatgc tccaggccctt ggcagactc 5280  
 acggccggaga cggtaacgc taggtgtggc agggatctc cccctctt ggcagctca 5340  
 tcagctagcc agctgtctgc gccttcctt aaggcaacat gcaactcccg tcatgactcc 5400  
 ccggacgtcg acctcatcg ggcacccctc ctgtggccgc aggagatggg cggaacatc 5460  
 acccgcgtgg agtcgaaaaa taaggtagta atttggact ctttcgagcc gctccaagcg 5520  
 gaggaggatg agagggaaatg atccgttccg gcggagatcc tgccggaggtc cagaaatttc 5580  
 cctcgagcga tgcccatatg ggcacggccg gattacaacc ctccactgtt agagtctgg 5640  
 aaggaccgg actacgtccc tccagtgtt caccgggtgtc cattggccg tgc当地 5700  
 cctccgatac caccctccacg gggaggagg acggttgc tgc当地 taccgtgtt 5760  
 tctgccttgg cggagctcgc agccgcacgg caacggccctt gccccccctt ccaatgc当地 5820  
 gtttagtcttgc actccctccat gggggggatcc aaggcaggag gcttagtgagg acgtcgatc 5880  
 gacgggtctt ggttaccgt tccatacacat ggacaggcgc cctgatcact cgtcgatc 5940  
 atcaatgcac tgaccaactc tttgtccgtt caccacaact acgtcgatc 6000  
 cgcagcgc当地 gcctgcggca gaagaaggtc acctttgaca cgggagatcc 5960  
 cactaccggg acgtgtctca gggatgttgg gcaaggcgatc 6060  
 ctatccgtgg aggaaggctg taagctgtt ccccccacatt cggcagatc taaatttggc 6120  
 tatggggcaa aggacgtcg cacaaggacc ttc当地 cccctt ggc当地 6180  
 tggaggact tgctggaaa cactgagaca ccaattgaca ccacatcat ggcaaaaaat 6240  
 gaggtttctt gctccaacc agagaagggg ggc当地 620  
 ccagatttgg ggggtcggtt gtgc当地 6300  
 cctcaggccg tgatgggtc ttccataact ttaaccacat ccgtccgtg 6360  
 ttccctgttga atgcctggaa agcgaaggaaa tggccstatgg ccaccatcat ggcaaaaaat 6420  
 tggggact caacgggtc tgagaatgtc atccgtgtt cagctcgctt taccatgt 6480  
 tggactgtt cccccc当地 gggatgttgg gcaaggcgatc 6540  
 gggggccccc tgacttaattc taaaggccg aactcggtgtt acgtgttgg ctccaccctc 6600  
 ggtgtactga cgaccagctg cggtaatacc ctacatgtt cttccgtt gccgggtcgag 6660  
 tggactgtc当地 cggagacga ccttgc当地 6720  
 atctgtgaaa gcgcccccc ccaagaggac gaggtcgatc tccaggccctt caccggaggct 6780  
 atgacttagat actctggccc ccctggggac cggccaaac atc当地 cccctt ggc当地 6840  
 ataacatcat gctccctccaa tggatgttgc ggc当地 6900  
 tatctcaccc gtgacccccc actccatgtt cgggagacga ccttgc当地 6960  
 actccatgtt gggggccctt ggc当地 7020  
 atgatccatga tgacttccat cttccatgtt cttccatgtt cgggagacga ccttgc当地 7080  
 cttagattgtc agatctacgg ggc当地 7140  
 attcaacgac tccatggccct tagcgc当地 tcaactccata cttccatgtt cgggagacga 7200  
 aatagggtgg cttcatgtcc caggaaactt ggggtaccgc cttgc当地 7260  
 cgggccc当地 gtgtccgatc taggtactgt tccagggggg gggggatcc 7320  
 aatagggtgg cttcatgtcc caggaaactt ggggtaccgc cttgc当地 7380  
 attcaacgac tccatggccct tagcgc当地 tcaactccata cttccatgtt cgggagacga 7440  
 aatagggtgg cttcatgtcc caggaaactt ggggtaccgc cttgc当地 7500  
 cgggccc当地 gtgtccgatc taggtactgt tccagggggg gggggatcc 7560

CA 02303526 2000-03-31

aagtacctct tcaactgggc agtaaggacc aagctcaaac tcactccaat cccggctgcg 7620  
tcccagggtgg atttataccag ctgggtcggt gctgggtaca gcgggggaga catatatcac 7680  
agcctgtctc gtgcccgacc ccgctgttc atgtggtgcc tactccact ttctgttaggg 7740  
gtaggcatct atctactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800  
ccatcctgtt tttttccctt ttttttttc tttttttttt tttttttttt 7860  
tttttcttctt tttttttctt cttttttcc tttttttcc tttgggtggct ccaccttagc 7920  
cctagtcacg gctagctgtg aaaggccgt gagccgcttg actgcagaga gtgctgatac 7980  
tggcctctct gcagatcaag t 8001

<210> 4  
<211> 7989

<212> DNA  
<213> Hepatitis C Virus

<400> 4

tataggctgg gagccgttca aaacgagggtt actaccacac acccccataac caaatacata 3660  
 atggcatgca tgcggctga cctggagggtc gtcacgagca cctgggtgt ggtaggcgga 3720  
 gtccttagcag ctctggccgc gtattgcctg acaacaggca gcgtggtcat tggggcagg 3780  
 atcatcttgt ccggaaagcc ggccatcatt cccgacaggaa aagtccstta ccggaggttc 3840  
 gatgagatgg aagaqgtgcgc ctcacaccc ctttatcatcg aacaggaaat gcaactcgcc 3900  
 gaacaattca aacagaaggc aatcggttg ctgcaaacag ccaccaagea agcggaggct 3960  
 gctgtccccg tgggtggat caagtgcgg accctcgaag cttctggc gaagcatatg 4020  
 tggaaatttca tcagcgggat acaatattta gcaaggcttgc ccactctggc tggcaacccc 4080  
 gcgatagcat cactgatggc attcacaccc tctatcacca gcccgcac caccaacat 4140  
 accctcttgt ttaacatcct gggggatgg gtggccgccc aacttgetcc tcccaagcgct 4200  
 gcttcgttgc tgcgtggcgc cgccatcgct ggacgggtc ttggcagcat aggcccttggg 4260  
 aagggtcttgc tggatatttt ggcagggttat ggacgggggg tggcaggcgcc gctcgtggcc 4320  
 tttaaggctca tgagcggcga gatgccctcc accgaggacc tggtaacct actccctgt 4380  
 atccctcccc ctggcggccct agtcgtcggt gtcgtgtcg cagcgatact gctcggcac 4440  
 gtggggccag gggagggggc tgcgtggg atgaacggc tgatagcggt cgcttcgggg 4500  
 ggttaaccacg tctcccccac gcaactatgtc cctgagagcg acgctgcagc acgtgtca 4560  
 cagatctct ctagtcttac catcaactcg ctgctgaaga ggcttcacca gtggatcaac 4620  
 gaggactgtc ccacgcccatt ctcgggtcg tggctaagag atgtttggg ttggatatgc 4680  
 acgggtttga ctgatttcaa gacctggcgc cagtcacaa ccctgcccgtt attgcccggg 4740  
 gtccctcttct tctcatgtca acgtgggtac aaggaggatct ggccggggcga cggcatcatg 4800  
 caaaccaccc gcccatgtgg agcacaatgc accggacatg tgaaaaacgg ttcctatggg 4860  
 atcggtgggc cttaggacatg tagtaacacg tggcatggaa cattcccccataacgcgtac 4920  
 accacggggcc ctcggcggcc cttccggcga ccaattttt cttagggcgt gtggcgggtg 4980  
 gctgtggagg agtacgtgg gttacgcgg gtgggggatt tccactacgt gacgggcgtg 5040  
 accactgaca acgtaaatgt cccgtgtcag gttccggccc ccgaattttt cacagaagtg 5100  
 gatgggggtgc ggttgcacag gtacgctcca gcgtgcaaac ccctcttacg ggaggagggtc 5160  
 acattcttgg tccggctcaa tcaatacccg gttgggtcac agtcccatg cgagcccgaa 5220  
 ccggacgttag cagtgctcac ttccatgcac accgaccctt cccacattac ggccggagacg 5280  
 gctaaggctgaa ggctggccag gggatcttcc ccttccttgg ccagctcactt agctagccag 5340  
 ctgtctgcgc cttcccttggaa ggcaacatgc actaccgtc atgactcccc ggacgctgac 5400  
 ctcatcgagg ccaaccttctt ctggcggcag gagatggggc ggaacatcac ccgcgtggag 5460  
 tcgaaaataa aggttagtaat tttggactt ttcgagccgc tccaaggcga ggaggatgag 5520  
 agggaaatgtt ccgttccggc ggagatctg cggaggtcca ggaaattttcc tcgagcgatg 5580  
 cccatatggg cacggccggaa ttacaaccctt ccactgttag agtcttgaa ggaccggac 5640  
 tacgtccctc ctaggttaca cgggtgttca ttggccgttgc ccaaggcccc tccgatacc 5700  
 cctccacggg ggaagaggac ggttgccttgc tcagaatcta ccgtgttcc tgccttggcg 5760  
 gagctcgccca caaagacattt cggcagctcc gaatcgctgg ccgtcgacag cggcacggca 5820  
 acggcccttc ctgaccggcc ctggcggac ggcgacgggg atctcagcga cgggttctgg 5880  
 tctccatgc ccccccatttga gggggagccg gggatcccc gtcgtatgtc tgcgtatgtt 5940  
 tctaccgtaa gcgaggaggc tagtggggac gtcgtctgtc agtgcgttccatccatgg 6000  
 acaggcgccc tgcgttccatccatgg tgcgtatgtc gaggaaacca agtgcgttccatccatgg 6060  
 agcaacttcc tgcgttccatccatgg ccacaacttgc gtctatgtca caacatctcg cagcgcaagc 6120  
 ctgcggcaga agaaggctac ctttgacaga ctgcagggtcc tggacgacca ctaccgggac 6180  
 gtgtctcaagg agatgaaggc gaaggcggtcc acagtttaagg ctaaacttct atccgtggag 6240  
 gaaggcctgtt agtgcgttcc cccacatcg gccagatcta aatttggctt tggggcaaaag 6300  
 gacgtccggg acctatcccg caaggccgtt aaccacatcc gctcgtgttgc gaggacttg 6360  
 ctggaaagaca ctgagacacc aatttgcaccc accatcatgg caaaaaaaaatgtt ggtttctgc 6420  
 gtccaaaccag agaaggggggcc cggcaagcca gctcgcctta tcgtattttcc agatttgggg 6480  
 gttcgtgtgt gcgagaaaat ggccctttac gatgtgggtct ccaccctcccc tcaggccgtg 6540  
 atgggctttt catacggtt ccaataactt cctggacagc gggtcgagtt cctggtaat 6600  
 gcctggaaag cgaagaaatg ccctatggc ttgcataatgc acacccgtt ttttactca 6660  
 acggtcactt agaatgcacat ccgtgttgc gaggtaatctt accaatgtt tgacttggcc 6720  
 cccgaaggcca gacaggccat aagggtcgcc acagagcgcc tttatcatgg gggcccccctg 6780  
 actaattctt aaggcgacaa ctgcggcttgc cgcgggtcc ggcgagccgg tgcgtatgtc 6840  
 accagctcgcc gtaataccctt cacaatgttac ttggggccgg ccgtcgccgt tgcgtatgtc 6900  
 aagctccagg actgcacat gtcgtatgc ggagacgacc ttgtcgatgtt ctgtgaaagc 6960  
 gccccggaccc aagaggacga ggcgagccca cggcccttca gggaggctat gactagatac 7020  
 tctggccccc ctggggaccc gcccaaaacca gaatacgact tggagggttgc aacatcatgc 7080  
 tcttccaatg tgcgtatgtc gcacgtatgc tctggcaaaa ggggttacta tctcaccctgt 7140  
 gaccccaacca ccccccatttgc gccccgttgc tggggacacag ttagacacac tccactgtt 7200  
 tcctggcttag gcaacatcat catgtatgtc cccacatttgc gggcaaggat gatccctgtat 7260  
 actcattttct tctccatccat tctgtatgtc gaaacatgtt aaaaaggccctt agattgtcag 7320  
 atctacgggg cctgttactc cattggggccatc ctggacttgc ctcagatcat tcaacgactc 7380  
 catggccatccat ggcgttccatccat gtcgtatgttacttgc gtcgtatgttacttgc 7440  
 tcatgcctca gggaaacttgg ggtaccggccc ttggcgttgc gtcgtatgttacttgc 7500  
 gtccgcgttca ggctactgtc ccagggggggg agggtgcacca ctgtggcaaa gtaccttttc 7560  
 aactggggcag taaggacaa gtcacaaacttgc actccaaatcc cggctcgatc ccagttggat 7620

CA 02303526 2000-03-31

ttatccagct gggtcggtgc tggttacagc gggggagaca tatatcacag cctgtctcg 7680  
gccccgacccc gctgggttcat gtggtgccct a ctccctacttt ctgttaggggt aggcatttat 7740  
ctactcccca accgatgaac ggggagctaa acactccagg ccaataaggcc atccctgtttt 7800  
tttccctttt tttttttctt tttttttttt tttttttttt ttctccctttt 7860  
tttttcctct ttttttcctt ttctttcctt tggtggtcc atcttagccc tagtcacggc 7920  
tagctgtgaa aggtccgtga gccgcttgac tgcagagagt gctgatactg gcctctctgc 7980  
agatcaagt 7989

<210> 5  
<211> 8649  
<212> DNA  
<213> Hepatitis C Virus

<400> 5  
gccagcccccc gattgggggc gacactccac catagatcac tccccctgtga ggaactactg 60  
tettcacgca gaaagcgtct agccatggcg ttagtatgag tgctgtcag cctccaggac 120  
ccccctcccc gggagagcca tagtggctg cgaaaccggg gagaatggg gattggccg 180  
gacgaccggg tccttttctt gataaeecc ctaatgcct ggagattgg gctggccccc 240  
gcfagactgs tagccgatg gtgttgggt gcggaaaaggcc ttgtgttact gcctgtatgg 300  
gtgcitgcga gtccccccggg aggtctcgta gaccgtgcac catgacacg aatctaaac 360  
ctcaaaagaaa aaccaaacgt aacacczaag ggcgcgcat gattgaacaa gatggattgc 420  
acgcagggttcc tccggccgct tgggtggaga ggctattcg gtagactgg gcacaacaga 480  
caatcggttgc ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccgggttctt 540  
ttgtcaagaa cgacctgtcc ggcccttga atgaactgca ggacgaggca ggcggccat 600  
cgtggctggc cacgacgggc gttcccttgcg cagctgtgtc cgacgttgc actgaagcgg 660  
gaagggactg gctgtattt ggcgaagtgc cggggcaggat tccctgtca tctcacctt 720  
ctctgcgca gaaatgttcc cggactctg cccatccgac cacaaggcga aacatcgat cgggtgtcat acgcttgatc 780  
tggaaaggccg tcttgtcgat caggatgatc tggacgaaaga gcatcagggg ctgcgcgcag 840  
ccgaaactgtt cgccagggtc aaggcgcga tgcccgcacgg cgaggatctc gtctgacccc 900  
atggcgatgc ctgttgcgg aatatcatgg tggaaaatgg ccgctttctt ggattcatcg 960  
actgtggccg gctgggtgtt ggccggcgt atcaggacat agcgttggct acccggtata 1020  
ttgtctgaaga gcttggcggc gaatgggctg accgcttctt cgtgttttac ggatccgcg 1080  
ctcccgatttgc gcaagcgtatc gccttctatc gccttcttga ctaggttgc ctagtttacat 1140  
cagaccacaa cgggttccctt ctagggat caattccgcg ccttccttcc cccccccctt 1200  
aacgttactg gccgaagccg tccacccat tgccgttctt cttggataaa ggccgggtgtt cttggatataa 1260  
acgagcatc cttaggggtct tggcaatgtg aggccccggg ttccccttc gccaaaggaa aacctggccc tgcgttcttgc 1320  
gtgaagggaaag cagttcttctt ggaagcttct tgaagacaaa tgcaagggtct tgcgttcttgc 1380  
tgcaggcagc ggaacccccc taagatacac ctgaaaggc aacccgttct acctggcgac aggtgcctt 1440  
gaaagagtca aatggcttc gtaaccctt gtcggatc tggcaatgtg ggccgggtgtt aacctggccc tgcgttcttgc 1500  
gtaccccttatt gtatgggatc tggatctggg cctcgggtgca gggcccccga accacggggg 1560  
tcgagggttaaaaacgtcta cggatataa ccatggaccg gggatggca gcatcggtcg cttggatataa 1620  
ctgatactct tgaccccttc tttacatatt ttatcaccag ggcggaggca cacttgcag aacgttactt 1680  
gttccgggggg gccgcgtgc ttaccatca caaaaatctt ctcacgttgc cgtcatctt cttggatataa 1740  
tttaccatca caaaaatctt gtcggccata ctccgttccac ctcacgttgc cgtcatctt cttggatataa 1800  
ataaccaaag tgccgtactt ctcacgttgc ctcacgttgc ctcacgttgc cgtcatctt cttggatataa 1860  
cggaagggtt ctgggggtca ggtacgtac tttatgacca gacccgtcg tggcgtgtt gtcacgttgc ctcacgttgc 1920  
tggggggcag acacccggc acgggggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 1980  
ctcgccctta ttacggctta agcctcacag gccgggacag acacaatctt ttcacgttgc ctcacgttgc ctcacgttgc 2040  
ggctcaaaaga ccctggccgg caggacactcg tcggctggca ggcagctcg acctttactt gtcacgtac tttatgacca 2100  
ggcagctcg acctttactt ggcacaca gggggggctt ggcgttccac tgctgtgtcc acctccacac gggggggctt gtcacgtac 2160  
ggcgttccac tgctgtgtcc acccgggggg ttcacgggtt ggtggacttt gtcacgttgc ctcacgttgc ctcacgttgc 2220  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 2280  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 2340  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 2400  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 2460  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 2520  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 2580  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 2640  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 2700  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 2760  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 2820  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 2880  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 2940  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 3000  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 3060  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 3120  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 3180  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 3240  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 3300  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 3360  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 3420  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 3480  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 3540  
ggcgttccac tgctgtgtcc acccgggggg agatacatct ttcacatatt ttatcaccag ggcggaggca cacttgcag 3600

aagctgtccg gcctcgact caatgctgta gcatattacc ggggccttga tgtatccgtc 3660  
 ataccaacta gcgagacgt cattgtcgta gcaacggacg ctctaattgac gggcttacc 3720  
 ggcgatttcg actcagtgtat cgactgcaat acatgtgtca cccagacagt cgacttcage 3780  
 ctggaccgcg ctttaccat tgagacgacg accgtgccac aagacgcggt gtcacgctcg 3840  
 cagcggcgag gcaggactgg tagggcagg atggcattt acagtttgt gactccagga 3900  
 gaacggccct cgggcatgtt cgatccctcg ttgtgtcgcg agtctatgac cgccggctgt 3960  
 gcttggatcg agtcacgccc cgccgagacc tcagtttaggt tgccggctta cctaaacaca 4020  
 ccagggttgc cctgtcgcca ggaccatctg gagttctggg agagcgtctt tacaggcctc 4080  
 aaccacatag acgcccattt cttgtcccg actaaggcagg caggagacaa cttccctac 4140  
 ctgttagcat accaggctac ggtgtcgcc aggctcagg ctccacctcc atcgtnnnac 4200  
 caaatgtgga agtgtctcat acggctaaag cctacgctgc acgggcaac gcccctgtc 4260  
 tataggctgg gagccgttca aaacgaggtt actaccacac accccataac caaatacatc 4320  
 atggcatca tgcgtcgctga cctggaggtc gtcacgacg cctgggtgtc ggttaggcgg 4380  
 gtcctagcg ctctggccgc gtattgcctg acaacaggca gctgtgtcat tggggcagg 4440  
 atcatcttgtt cggaaagcc ggcattt cccacaggg aagtcttta cccgaggttc 4500  
 gatgagatgg aagagtgccg aatcggttg ctgaaaacag aacagggaaat gcaactcgcc 4560  
 gaacaattca aacagaaggc aactgggtt cttacatcg ccaccaagca aecggaggct 4620  
 gctgctccc tggttggaaatc caagtggcg accctcgaag cttctggc gaagcatatg 4680  
 tggaaattca tcagcggtat acaatattt gcaaggcttgc 4740  
 gcgatagcat cactgtatggc attcacatcc tctatccca ccactctggc tggcaacccc 4800  
 accctctgtt ttaacatctt gggggatgg gtggccggcc aacttgccttcc tccacgcgt 4860  
 gttctgtt tcgttaggcgc cggcatcgct ggacggcgctg ttggcagcat aggcttggg 4920  
 aagggtctt tggttatttt ggcaggttat ggacgaggggg tggcggc 4980  
 tttaaaggta tgagcggtca gatggcccttcc accgaggacc tggtaacact actccctgt 5040  
 atcctctccc ctggccctt gttttttttt gtcgtgtcg tgatagcggtt cgcttcgccc 5100  
 gtggggccag gggagggggc ggttacccatc gcaactatgtt aacgtcgacg acgtgtcact 5160  
 ggttacccatc gttttttttt gtcgtgtcgcatgg ggttacccatc gtggatcaac 5220  
 cagatccctt ctatgttttcc gtcgtgtcgcatgg ggttacccatc atggttgggatggat 5280  
 gaggactgtt ccacggccatc gtcgtgtcgcatgg ggttacccatc cttccggctcg 5340  
 acgggtgttca ctgttccatc gtcgtgtcgcatgg ggttacccatc cttccggctcg 5400  
 gtcccttctt ttcgtatgtca gtcgtgtcgcatgg ggttacccatc ttcactacgt 5460  
 caaaccacatc gcccattgtgg gtcgtgtcgcatgg tggggccatc ttcgtatgtca 5520  
 atcggtgggc cttaggacatc gtcgtgtcgcatgg ggttacccatc ttcgtatgtca 5580  
 accacggggcc cctgcacgc gtcgtgtcgcatgg ggttacccatc ttcgtatgtca 5640  
 gtcgtgttca agtacgttca gtcgtgtcgcatgg ggttacccatc ttcactacgt 5700  
 accactgaca acgtttttttt gtcgtgtcgcatgg ggttacccatc ccgaattttt 5760  
 gatgggggttgc ggttgcacatc gtcgtgtcgcatgg ggttacccatc cttcccttccatc 5820  
 acatccctgg ttcggcttca gtcgtgtcgcatgg ggttacccatc agtcccttccatc 5880  
 ccggacgttccatc gtcgtgtcgcatgg ggttacccatc cttccatgttc accgaccctt 5940  
 gtcgtgttca ggttacccatc gtcgtgtcgcatgg ggttacccatc ccagcttccatc agtactcccc 6000  
 ctgtctggcc cttcccttgg gtcgtgtcgcatgg ggttacccatc atgactcccc 6060  
 ctcatcgagg ccaacccatc gtcgtgtcgcatgg ggttacccatc ggaacatcac 6120  
 tcagaaaata agttagtaat gtcgtgtcgcatgg ggttacccatc ttcactacgt 6180  
 agggaaatgtt ccgttccggc cccatatggg caccggccatc gtcgtgtcgcatgg 6240  
 tacgttccctt ctcgtgttca gtcgtgtcgcatgg ggttacccatc agtccctggaa 6300  
 cttccacatc gtcgtgtcgcatgg ggttacccatc ccaaggcccc tccgataccatc 6360  
 ggttgcgttccatc gtcgtgtcgcatgg ggttacccatc ccgtgttccatc tgccctggc 6420  
 gagctcgcca caaagacatc gtcgtgtcgcatgg ggttacccatc ccgtcgacatc 6480  
 acggcccttc ctgaccatc gtcgtgtcgcatgg ggttacccatc gatccgacatc 6540  
 tcctccatgc ccccccatttca gtcgtgtcgcatgg ggttacccatc atctcagatc 6600  
 tctaccatgc gtcgtgtcgcatgg ggttacccatc gtcgtgtcgcatgg 6660  
 acaggccccc tgcgtgtcgcatgg ggttacccatc atgcgtgtcgcatgg 6720  
 agcaacttcc tgcgtgtcgcatgg ggttacccatc ccacaacttcc gtcgtgtcgcatgg 6780  
 ctggccgaga agaaggatc gtcgtgtcgcatgg ttttgcacatc gtcgtgtcgcatgg 6840  
 gtgtcaagg agatgttcaatc gtcgtgtcgcatgg ggttacccatc ctaaaacttcc 6900  
 gaaaggcttccatc gtcgtgtcgcatgg ggttacccatc aatttggatc tggggcaaaat 6960  
 gacgtccggatc acctatccatc gtcgtgtcgcatgg ggttacccatc caaggccgtt 7020  
 ctggaaagaca ctgacatc gtcgtgtcgcatgg ggttacccatc aatttggatc ggttttctgc 7080  
 gtccaaaccatc agaagggggg gttcgtgttgc gtcgtgtcgcatgg ggttacccatc cccatatggc 7140  
 atgggtctt catacgatc gtcgtgtcgcatgg ggttacccatc ttcggccgtt 7200  
 gcctggaaatc cgaaggaaat gtcgtgtcgcatgg ggttacccatc ggttgcgtt 7260  
 acggtcactt gtcgtgtcgcatgg ggttacccatc cccatatggc ttcggccgtt 7320  
 cccgaaggccatc gacaggccatc aaggatcgttccatc gtcgtgtcgcatgg 7380  
 actaatttca aaggccatc gtcgtgtcgcatgg aaggatcgttccatc acaggccgtt 7440  
 accagcttccatc gtaataccatc ccatgttac ttgttggccatc gtcgtgtcgcatgg 7500  
 aagctccagg actgcacatc gtcgtgtcgcatgg ggttacccatc ggttgcgtt 7560  
 aagctccagg actgcacatc gtcgtgtcgcatgg ggttacccatc ggttgcgtt 7620

gcggggaccc aagaggacga ggcgagccta cgggcctca cgaggctat gactagatac 7680  
tctccccccc ctggggaccc gcccaaacca gaatacgact tggagtttat aacatcatgc 7740  
tcctccaatg tgtcagtgcgc gcacgatgca tctggcaaaa gggtgtacta tctcaccgt 7800  
gaccccacca ccccccttgc gcgggctgcg tggagacag ctagacacac tccagtcaat 7860  
tcctggctag gcaacatcat catgtatgcg cccaccttgc gggcaaggat gatcctgtatg 7920  
actcattct tctccatctc tctagctcag gaacaacttg aaaaaggctt agattgtcag 7980  
atctacgggg cctgttaactc cattgagcca cttgacacctac ctcagatcat tcaacgactc 8040  
catggcccta ggcattttc actccatagt tacttccag gtgagatcaa taggtggct 8100  
tcatgcctca ggaacttgg ggtaccggcc ttggagatct ggagacatcg ggcagaagt 8160  
gtcccgctca ggctactgtc ccaggggggg agggtgccaa ctgtggcaa gtacctttc 8220  
aactgggcag taaggaccaa gctcaaactc actccaatcc cggctgcgtc ccagttggat 8280  
ttatccagct ggttcgttgc tggttacagc ggggagaca tatatcacag cctgtctcg 8340  
gcccggaccc gctgggtcat gtggtgctca ctctacttt ctgttaggggt aggcatctat 8400  
ctactcccca accgatgaac ggggagctaa acactccagg ccaataggcc atctgtttt 8460  
tttccctttt tttttttttt tttttttttt tttttttttt ttctccctttt 8520  
tttttcctct ttttttcctt ttctttcctt tggtggctcc atcttagccc tagtcacggc 8580  
tagctgtgaa aggtccgtga gccgcttgac tgcagagagt gctgatactg gcctctgtc 8640  
agatcaagt 8649

<210> 6  
<211> 8001  
<212> DNA  
<213> Hepatitis C Virus

<400> 6

```

gccagcccc gattggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacga gaaagcgct agccatggcg ttagtatgag tgtcgtcag cctccaggac 120
ccccctccc gggagagcca tagtggctg cgaaaccgg gatcacaccg gaattgccag 180
gacgaccggg tccttcttg gatcaacccg ctaatgcct ggagatttg gcgtgcccc 240
gcgagactgc taaccgagta gtgttggtc gcaaaggcc ttgtggact gcctgatagg 300
tgtctgcga gtccccccggg aggtctcgta gaccgtc acgacacg aatccaaac 360
ctcaaagaaa aaccaaacgt aacaccaacg ggccgcacat gattgaaaca gatggattgc 420
acgcagggtc tccggcgct tgggtggaga ggctattcg ctatgactgg gcacaacaga 480
caatcgctg ctctgtatgcc gccgtttcc ggtgtcagc gcaggggcgc ccggttctt 540
ttgtcaagac cgaccgttcc ggtgcctgaa atgaactgca ggacgaggca ggcggctat 600
cgtggctggc cacacggggc gtccctgtcg cagctgtgct cgacgttgc actgaagcgg 660
gaaggggactg gtcgttattg ggcaagtgc cggggcagga tctctgtca tctcacctt 720
tcctgcga gaaagtatcc atcatggctg atgcaatgcg gcccgtcat acgcttgatc 780
cggttacctg cccattcgac caccaagcga aacatgcac gtaggttgcgca 840
tggaaagccgg tcttgcgtat caggatgatc tggacgaaga gcatcagggg ctcgcggccag 900
ccgaacgtt cgccagggtc aaggcgcgc tgccgcacgg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgttgcgg aatatcgatgg tggaaaatgg cccgtttctt ggattcatcg 1020
actgtggccg gctgggtgtg gggaccgt atcaggacat acgttggct acccggtata 1080
ttgtcaaga gcttggccg gatgggctg accgttctt cgtgtttac ggtatcgccg 1140
tcctcgatcc gcaacgcac gcttctatc gccttcttgc ctagtgcggat caattccgc cctctccctc cccccccct 1260
cagaccacaa cggttccct aacgttactg gccgaataa ggccgtgtg cttttgtcta tatgttattt 1320
tccaccat tgcgtctt tggcaatgtg aggccccggg aacctggccc tgcgttctt 1380
acgagcattc ctggggctc ttccctctc gccaaaggaa tgcaagttct gttaatgtc 1440
gtgaaaggaa cagttcttctt ggaagcttct tgaagacaaa caacgtctgt acgcaccctt 1500
tgcaggcgc ggaacccccc acctggcgc aggtgcctct gggccaaaa gccacgtgt 1560
taagatacac ctgaaaggc ggcacaaccc cagtgcacg ttgtgagttt gatagttgt 1620
gaaagagtca aatggcttc ctcaagcgta ttcaacaagg ggctgaagga tgcccaagaag 1680
gtacccctt gtatggatc tgatctggg cctcgggtca catgtttac atgtgtttag 1740
tcgagggttaaaaacgtcta ggccccccga accacggggg cgtgttttc cttgaaaaa 1800
cacgataata ccatggcgcc tattacggcc tactccaaac agacgcgagg cctacttggc 1860
tgcacatca ctggcctac agggccggc aggaacaggc tgcaggggga ggtccaagtg 1920
gttccaccg caacacaatc ttcctggcg acctgcgtca atggcgtgt ttggactgtc 1980
tatcatggc cccgtctaa gacccttgc gggccaaagg gccaatc accaaatgtac 2040
accaatgtgg accaggacat cgtggctgg caagcgc cccggggcg ttccttgaca 2100
ccatgcaccc tggcgacgtc ggaccttac ttggtcacga ggcattccg tgcattccg 2160
gtgcggccgc gggcgacag cagggggcgt ctactctcc ccaggccgt ctctactt 2220
aagggtctt cggcggtcc actgtctgc ccctcgggc atgtctggg catcttccg 2280
gctggcggt gcaacccgagg gtttgcgaaq ggggtggact ttgtaccctg cgagtctatg 2340
gaaaccacta tgggtcccc ggttccacg gacaactctg cccctccggc ctagccgc 2400
acatccagg tggcccatc acacgcctt actggtagcg gcaagagcac taaggtccg 2460
gctgcgtatc cagcccaagg gtataagggt ctgttctgaa acccgccgt cggccacc 2520
ctaggttcg gggcgatatt gtctaaaggca catggtacgc accctaacat cagaaccggg 2580
gtagggacca tcaccacggg tggcccatc acgtacttca cctatggcaa gtttcttgcc 2640
gacgggtgtt gctctggggg cgcctatgac atcataatat gtgtatggatg ccactcaact 2700
gactcgacca ctatcttggg catcgacca gtctggacc aagcgagac ggctggagcg 2760
cgactcgctg tgctcgccac cgctacgcct cgggatcggt tcaccgtcc acatccaaac 2820
atcgaggagg tggctctgtc cagactggaa gaaatcccc tttatggcaa agccatcccc 2880
atcgcgacca tcaagggggg gagggcaccctt atttcttgc atttcaagaa gaaatgtat 2940
gagtcgcggc cgaagcttcc cggcctcgga ctcaatgtg tagcatatta cccggggctt 3000
gatgtatccg tcataccaac tagcgagac gtcatgtcg tagcaacggc cgctctaatt 3060
acgggttta cccgcgattt cgactcgatc atcgactgca atacatgtt caccacgaca 3120
gtcgacttca gcctggaccc gacccatcatttgcgaa cggccgtcc acaagacgcg 3180
gtgtcaatcg cgcaggccg aggccggact ggttagggca ggatggcat ttacaggttt 3240
gtgactcccg gagaacggcc ctggccatg ttcatgttcc cgggttctgt cgagtgttat 3300
gacgcgggct gtgttggta cgactcgatc cccggccgaga cctcaggtag ttgcgggct 3360
tacctaaca caccagggtt cccgttgc caggaccatc tggatgtcg ggagagcgtc 3420
tttacaggcc tcaaccacat agacgcctt ttcttgc tccatgttcc agactaggca ggaggagac 3480
aacttccctt acctggtagc ataccaggct acgggtgtcg ccagggtctc ggctccaccc 3540
ccatcgccgg accaaatgtg ggagtgtctc atacggctaa agcctacgct gacccggcc 3600

```



tcccagttgg atttatccag ctgggtcggt gctgggttaca gcgggggaga catatatcac 7680  
agcctgtctc gtccccgacc ccgcgtggtc atgtgggtgcc tactcctact ttctgttaggg 7740  
gtaggcatct atctactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800  
ccatcctgtt tttttccctt ttttttttcc tttttttttt tttttttttt tttttttttt 7860  
ttttctcctt tttttttccctt ctttttttcc ttttttttcc tttgggtggct ccatcttagc 7920  
ccttagtcacg gctagctgtg aaaggtccgt gagccgcttg actgcagaga gtgctgatac 7980  
tggcctctct gcagatcaag t 8001

<210> 7  
<211> 11076  
<212> DNA  
<213> Hepatitis C Virus

<400> 7

```

gcccggccccc gattgggggc gacactccac catagatcac tccccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgcgtgcag cctccaggac 120
ccccccctccc gggagagcca tagtggtctg cggAACCGGT gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaaccsg ctcaatgcct ggagatttg gcgtgcggcc 240
gcgagactgc tagccgagta gtgttgggtc gcggAAAGGCC ttgtggact gcctgtatagg 300
gtgcttgcga gtggcccccggg aggtctcgta gaccgtgcac catgagcacg aatccctaaac 360
ctcaaagaaa aaccaaacgt aacaccaacg ggcgcgccc gattgaacaa gatggattgc 420
acgcagggtt tccggccgt tgggtggaga ggctattcg tcatgactgg gcacaacaga 480
caatcggtc ctctgtatcc gcccgttcc ggctgtcagc gcaggggcgc ccgggttctt 540
ttgtcaagac cgacctgtcc ggtgcctgtc atgaactgcg ggacgaggca ggcggctat 600
cgtggctggc cacgacgggc gttccttgcg cagctgtgtc cgacgttgtc actgaagcgg 660
gaagggactg gctgtatgg ggcgaagtgc cggggcagga tctcctgtca tctcacctt 720
ctsctgcga gaaagtatcc atcatggctg atgaatgcg ggcggctgtcat acgcgttgc 780
cggtacactg cccattcgac caccaagcga aacatgcat cgagcgagca cgtaactcgga 840
tggaaagccgg tcttgcgtat caggatgate tggacgaga gcatcagggg ctcgcggcag 900
ccgaactgtt cgccagggtc aaggcgogca tgcccgcacgg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgttgcgg actgtggccg gctgggtgtg gatcgatgcg ggcgtttctt ggattcatcg 1020
ttgtctgaaga gcttggccgc gaatgggtg accgcttcc gatcgatgcg ggcgtttctt 1080
ctcccgttcc gcagcgcattc ctaggttccct aatcatggggat caattccgc cctctccctt ccccccctt 1140
cagaccacaa cggttccctt aacggtactg gccgaaggccg atcaggacat ggcgttgc 1200
tccaccatg tggcgcttt acggatcc cttaggggtct gtaaggaaag caggctctt gcaaggctt 1260
gtgaaggaaat gtcaggcagc ggaacccccc taagatacac ctgcaaaaggc ggcacaaccc ctttgcgtatc tttttttttt 1320
ggcacaaccc ctttgcgtatc tggcaatgtg aggccccggg aacctggccc tttttttttt 1380
acggatcc cttaggggtct tttttttttt gcaaggctt 1440
gttgcgtatc gtcaggcagc ggaacccccc taatccatc ccccttccatc ccccttccatc 1500
acgttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 1560
ttttttttt gcaaggctt 1620
ggcacaaccc ctttgcgtatc tttttttttt 1680
ctcaagcgtt tttttttttt 1740
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 1800
ggcacaaccc ctttgcgtatc tttttttttt 1860
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 1920
ttttttttt 1980
caacctcgatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 2040
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 2100
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 2160
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 2220
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 2280
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 2340
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 2400
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 2460
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 2520
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 2580
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 2640
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 2700
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 2760
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 2820
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 2880
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 2940
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 3000
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 3060
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 3120
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 3180
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 3240
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 3300
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 3360
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 3420
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 3480
gttgcgtatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 3540
aaacacccatc gtcaggcagc ggcacaaccc ctttgcgtatc tttttttttt 3600

```

tgggttcgg ggcccttgggtt gacaccaggaa tgcttggtcc actaccata caggctttgg 3660  
 cactaccctt gcaactgtcaa ctttaccatc ttcaaggtaa ggatgtacgt ggggggagtg 3720  
 gagcacaggc tcgaaggccgc atgcaattgg actcgaggag akgcgttgc aa cctggaggac 3780  
 agggacagat cagagcttag ccgcgtctg ctgtctacaa cggagtggca ggtattgccc 3840  
 tgttccctca ccaccctacc ggctcgatc actggttga tccatctcca tcagaacgtc 3900  
 gtggacgtac aataacctgtaa cgttataggg tcggcggttg tctctttgc aatcaaattgg 3960  
 gatgtatgtcc tggctgtctt ccttcttctg gcgacgcgc gctgtctgtc ctgtttgtgg 4020  
 atgatgtgc tgatagctca agctgaggcc cccttagaga acctgggtt cctcaacgcg 4080  
 gcatccgtgg ccggggcgca tggcatttctc tccttctcg tggctttctg tgctgcttgg 4140  
 tacatcaagg gcaaggctggc ccctggggcg gcatatgccc tctacggcgt atggccgcta 4200  
 ctctctgtcc tggctggcggtt accaccacga gcatacgcga tggaccggga gatggcagca 4260  
 tcgtgceggag ggcgggtttt cgttaggtctg atactttaa ccttgcgtcc gcaactataag 4320  
 ctgttccctcg cttaggtcat atgggttta caatattttt tcaccaggc cgaggcacac 4380  
 ttgcaagttt ggatcccccc cctcaacgtt cggggggggcc gcatgtccgt catctcttc 4440  
 acgtgcgcga tccaccaggaa gctaaatctt accatcaacca aaatcttgcg cccataactc 4500  
 ggtccactca tggctgtccca ggctggata accaaaatgtc cgtacttcgt gcgcgcacac 4560  
 gggctcattc gtgcattgcatt gctggcgcgg aaggttgcgt ggggtcatta tggccaaatg 4620  
 gctctcatga agttggccgc actgacaggt acgtacgtt atgaccatct caccctactg 4680  
 cgggacttggg cccacgcggg cctacgagac ctggcggtgg cagttgagcc cgtcgcttc 4740  
 tctgatatgg agaccaaggta tattacccgtt gggcagaca cccggcggtg tggggacatc 4800  
 atcttggggcc tggccgtctc cggccggcgg gggaggaga tatacttggg accggcagac 4860  
 agccttgaag ggcgggggtg gcaactctc ggccttata cggcctactc ccaacagacg 4920  
 cgaggcttac tggctgtcat catcaactagc cttcacaggcc gggacaggaa ccaggtcgag 4980  
 ggggagggtcc aagtggtctc caccgcacca caatcttcc tggcgacctg cgtcaatggc 5040  
 gtgtgttggaa ctgtctatca tggtgccggc tcaaaagaccc ttggccggccc aaaggggccca 5100  
 atcaccctaaa tgtacaccatc tgggaccatc gacccgtctg gctggcaagc gccccccggg 5160  
 ggcgcgttccct tgacaccatc caccctgcggc agtccggacc ttacttggg caccaggcat 5220  
 gccgatgtca ttcccggtgcg cccggggggc gacagcaggg ggagcctact ctccttccagg 5280  
 cccgttctact acttgaaggta ctttccggc ggttcaactgc tcttccctc gggcatgtc 5340  
 gtggccatct ttcgggtctc cgtgtgcacc cgggggggtt cgaaggccgt ggactttgtt 5400  
 cccgtcgagt ctatggaaac castatgcgg tcccccgtct tcacggacaa ctcgtccct 5460  
 cccggcttac cgcagacatt ccagggtggc catctacacg cccctactgg tagcggcaag 5520  
 agcaactaagg tggccgtctc gtatgcagcc caagggtata aggtgtttt cctgaacccg 5580  
 tccgtcgccg ccaccctagg tttccggcg tataatgtcta aggacatgg tatcgacact 5640  
 aacatcagaa ccggggtagg gaccatcacc acgggtgcgg ccatcacgtt ctcacccat 5700  
 ggcaagttt tcggccacgg tgggtgtctt gggggccct atgacatcat aatatgtgtat 5760  
 gagtgcact caactgactc gaccactat ctgggcatcg gacacgtctt ggaccaagcg 5820  
 gagacggctg gaggcgact cgtgtgtctc gccacccgtt cgcctccggg atcggcattc 5880  
 gtgccacatc caaacatcga ggagggtggct ctgtccagca ctggagaaat ccccttttat 5940  
 ggcaaagcca tcccccattc gaccatcaag gggggggaggc acctcattt ctggccattcc 6000  
 aagaagaaat gtatggatct cggccgcaag ctatccggcc tcggactcaa tgctgttagca 6060  
 tattaccggg gccttgtatc atccgttata ccaacttagcg gagacgtcat tgctgttagca 6120  
 acggacgctc taatgacggg ctttaccggc gatccgtact cagtgtatcg ctgcaatata 6180  
 tgggtcaccct agacagtctc ctttccggc gaccggaccc tcaccatgtt gacgacgacc 6240  
 gtgccacaaag acgggtgtc acgtccgtc cggcgaggca ggactggtag gggcaggatg 6300  
 ggcattttaca ggttgggtac tccaggagaa cggccctcg gcatgttgc ttctcggtt 6360  
 ctgtgcgtgt gctatgtacgc gggctgtctt tggtacgacg tcacggccgc cgagacctca 6420  
 gttaggttgc gggcttaccc aaacacacca gggctggcccg tctggcagga ccatctggag 6480  
 ttctgggaga gcttctttac aggccttacc cccatagacg cccattttt gtcccagact 6540  
 aggcaggcag gagacaactt cccctacctg gtacgatacc aggttacgggt gtgcggccagg 6600  
 gctcaggctc cacccatcgtt gtgggaccaa atgtgggtt gtcctatcgt gctaaagct 6660  
 acgtcgacg ggcacacggc cctgtgtat aggtggggat cgggttactt cgggttact 6720  
 accacacacc cccatataaa atacatcatc gcatgtatc cggctgaccc ggagggtcg 6780  
 acggacactt ggggtgtgtt aggcggggatc ctggcgttc tggccgcgtt ttgcctgaca 6840  
 acaggcagcg tggcttcatgt gggcaggatc atcttgcgtt gaaaggccgc catcattecc 6900  
 gacaggaaag tccttaccg ggagttcgat gagatggaaat gtgcgcctc acacccctt 6960  
 tacatcgaac agggaaatgc gtcggccaa caattcaaa acggacat cgggttgcgt 7020  
 caaacagcca ccaaggcaagc ggaggctgtt gctccctgtt tggaaatccaa gtggccggacc 7080  
 ctcgaaggctt tctggccgaa gcatatgtgg aatttcatca gcccggatata atatttagca 7140  
 ggcttgcgttca ctctgcgttgg caaccccgcc atagcatcgt tggatggcatt cacaggctt 7200  
 atcaccaggc cgttcaccatc ccaacatcatc ctctgtttt acatcttggg gggatgggtt 7260  
 gccggccaaac ttgtcttcc cagcgctgtt tggcttgcgtt tagggccggg catcgcttgc 7320  
 gcccgtgtt gcaactatgg ctttggggaaat gtcgttgcgtt atattttggc aggttatgg 7380  
 gcaagggtgg caggcgccgtt cgtggccctt aaggatcatca gcccggat gcccctccacc 7440  
 gaggacactgg ttaacacttcc cctgtatc ctctccctt gccccttgcg cgtcgccggc 7500  
 gtgtgcgtcag cgataactgcg tcggcactgtt ggcccgagggg agggggctgt gcaactggat 7560  
 aaccggctga tagcgatcgatc ttcgggggtt aaccacgtt ccccccacgcg cttatgtgtt 7620

gagagcgcacg ctgcagcacg tgcactca gatcttccat gtcttaccat cactcagctg 7680  
 ctgaagaggc ttccccatcg gatcaacgag gactgctcca cgccatgtc cggtctgtgg 7740  
 ctaagagatg tttggattg gatatgcacg gtgttactg atttcaagac ctggctccag 7800  
 tccaagctcc tgccgcgatc gcccggagtc cccttcttcatgtcaacg tgggtacaag 7860  
 ggagtctggc gggcgcacgg catcatgcaa accacctgcc catgtggac acagatcacc 7920  
 ggacatgtiga aaaacggttc catgaggatc gtggggccta ggacctgttag taacacgtgg 7980  
 catggAACAT tccccatcaa cgcgtacacc acggggccct gcacgcctc cccggcgcca 8040  
 aatttattca gggcgctgtg gccccggct gctggggact acgtggatgt aacttccat 8100  
 ggggattcc actacgtgac gggcatgacc actgacaacg taaagtgcggc gtgtcagggt 8160  
 cccggcccc aattttcac agaagtggat ggggtgcggc tgcacaggta cgctccagcg 8220  
 tgcaaacccc tcctacggga ggaggtcaca ttccctggtc ggctcaatca atacctgttt 8280  
 gggtcacagc tcccatgcgg gcccgaaccg gacgttagcag tgctcaatcc catgctcacc 8340  
 gaccctccc acattacggc ggagacggct aagcttaggc tggccagggg atctccccc 8400  
 tccttggcca gctcatcagc tagccagctg tctgcgcctt ctttgaaggc aacatgcact 8460  
 acccgtcatg actcccccgg acatcccccgg cgtggatca gaaaataagg tagtaatttt ggactctttc 8580  
 gaggcgcctc aaggcgagg gatgagggaa gaatgtatcc ttccggcggaa gatcctgcgg 8640  
 aggtccagga aattccctcg agcgatgccc atatggcact gcccggatta caaccctcca 8700  
 ctgttagagt ccttggagga cccggactac gtccctccag tggtacacgg gtgtccattg 8760  
 ccgcctgcca aggccccctcc gataccaccc caacggagg agaggacgggt tgcctgtca 8820  
 gaatctaccg tgccttctgc cttggcgag ctcgcaccaa agaccttcgg cagctccgaa 8880  
 tcgtcgcccg tcgacagcgg cacggcaacg gctcttcgt accagccctc cgacgacggc 8940  
 gacgcggat ccgacgttgatcgactcc tccatgtcccccc cccttgggggg 9000  
 gatccccgtc tcaagcggcgtt gtcgggttcc accttaaggc aggggttag tgaggacgtc 9060  
 gtctgctgtc cgatgtccta cacatggaca ggcgcctga tcaacccatcg tccgtcacca caacttggtc 9180  
 gaaaccaagc tgcccgtaa tgcaactgagc aactcttgc aggtcacctt tgacagactg 9240  
 tatgtaccaa catctcgacg cgcaaggctg cggcagaaga tgaaggcggaa ggcgtccaca 9300  
 caggccctgg acgaccacta ccgggacgtg ctcaaggaga gtttgggactt ctttacgtatcg 9360  
 gtttgggactt aacttctatc cgtggaggaa gctgtaaagc tatccagcaa ggcgttaac 9420  
 agatctaat ttggctatgg ggcggaaaggac gtcggaaacc agacaccaat tgacaccacc 9480  
 cacatccgcg cctgtggaa ggacttgcg gaagacactg atgacatccg tgggggggg 9540  
 atcatggcaa aaaaatgggt tttctgtc caaccaggaga gggccataag gtcgctcaca 9840  
 cgccttatcg tattccaga gtttgggggtt cgtgtgtgcg agaaaatggc ctttacgtatcg 9600  
 gtggcttcca ccctccctca ggccgtgatc ggctcttcat acggattcca atactctcct 9660  
 ggacagcggg tcgagttctt ggtgaatgcc tggaaagcga agaaatggcc tatggcttc 9720  
 gcatatgaca cccgtgttt tgactcaacg gtcactgaga atgacatccg tgggggggg 9780  
 tcaatctacc aatgttgtga ctggcccccc gaagccagac ggcagaactg cgctatcgc 9900  
 gagcggctt acatcgggggg cccctgtact aattctaaag ataccctcact atgttatttg 9960  
 cgggtccgcg cgagcgggtgt actgacgacc agtgcggta agtgcgtatccatcg 10020  
 aaggccctgtc cggccctgtcg agtgcgaag ctccaggact gacatgtatccatcg 10080  
 gacgacccctt tcgttatctg tgaaagcggc gggacccaag gggaccggcc caaaccagaa 10140  
 gccttcacgg aggtatgac tagatactt gccccccctg cgtcgccatcg 10200  
 tacgacttgg agttgataac atcatgtcc tccaaatgtgt cccatcgatct 10260  
 ggcaaaagggt tgcactatctt caccgtgac cccaccaccc cccttgcgcg ggctgcgtgg 10320  
 gagacagcta gacacactcc agtcaattcc tggcttaggc acatcatcat gtatgcgc 10380  
 accttgggg caaggatgtatctt cctgtatactt catttcttccatcgatct 10440  
 caacttggaa aagccctaga ttgtcagatc taccggggctt ctttccatcg 10500  
 gacacttcc agatcatca agactccat ggccttagcg ctttccatcgatct 10560  
 tctccaggatgt agatcaatag ggtggcttca tgcctcaggaaacttgggggtt accggcccttgc 10620  
 cgagtctggc gacatcgggc cagaagtgtc cgccgttaggc tactgtccca ggggggggg 10680  
 gctgccactt gtggcaagta cctttcaac tggcagtaa ggaccaagct caaactcact 10740  
 ccaatcccg ctgcgtccca gttggattta tccagctggc tgcgtgtgg ttacagcggg 10800  
 ggagacatata tcacacgcct gtctcgatcg cggcccttgc gtttcatgtg gtgcctactc 10860  
 ctacttctcg tagggtagg catctatcta ctcccccaacc gatgaacggg gagctaaaca 10920  
 ctccaggcca ataggccatc ctgtttttt ccctttttt ttttctttt tttctttt tttctttt 10980  
 tttttttt tttttttt tttttttt tttttttt tttctttt tttctttt tttctttt 11040  
 tggctccatc ttagccctag tcaaggcttag ctgtgaaagg tccgtgagcc gcttgactgc 11076  
 agagagtgtcaacttgcagca tcaagt

<210> 8  
<211> 8001  
<212> DNA  
<213> Hepatitis C Virus

<400> 8

gcacggccccc	gattgggggc	gacactccac	catacatcac	tcccctgtga	ggaactactg	60
tcttcacgca	gaaagcgtct	agccatggcg	ttatgtatgag	tgtcggtcag	cctccaggac	120
ccccccctccc	gggagagcca	tagtggctcg	cggaaaccgg	gagtagccac	gaattgccag	180
gacgaccggg	tcccttcttg	gatcaacccg	ctcaatgcct	ggagatttgg	gcgtgcccc	240
gcgagactgc	tagccgagta	gtgttgggtc	cgaaaaggcc	ttgtggtaact	gcctgatagg	300
gtgcttgcga	gtgcccggg	aggtctcgta	gaccgtgcac	catgagcacg	aatcctaaac	360
ctcaaagaaa	aaccaaacgt	aacaccaacg	ggcgcgcac	gattgaacaa	atggattgc	420
acgcagggtc	tccggccgt	tgggtggaga	ggcttattcg	ctatgactgg	gcacaacaga	480
caatcggtc	ctctgtatcc	gcccgttcc	ggctgtcagc	gcaggggcgc	ccggttctt	540
ttgtcaagac	cgacactgtcc	ggtgccttga	atgaactgca	ggacgaggca	gcgcggctat	600
cgtggctggc	cacgacgggc	gttccttgc	cagctgtgt	cgacgttgc	actgaagcgg	660
gaaggggactg	gctgtatgg	ggcgaagtgc	cggggcagga	tctcctgtca	tctcacctt	720
ctccctgcga	gaaagtatcc	atcatggctg	atgcaatgcg	gcggctgcac	acgcttgatc	780
cggctacctg	cccattcgac	caccaaggca	aacatcgcat	cgagcgagca	cgtactcgga	840
tggaaagccgg	tcttgtcgat	caggatgatc	tggacgaaga	gcatcagggg	ctcgcgcag	900
ccgaactgtt	cgcagggtc	aaggcgcga	tgcggcagcc	cgaggatctc	gtcggtaccc	960
atggcgatgc	ctgttgcgg	aatatcatgg	tggaaaatgg	ccgttttct	ggattcatcg	1020
actgtggccg	gctgggtgt	gcggaccgt	atcaggacat	agcgttgct	accctgtata	1080
ttgctgaaga	gcttggccgg	gaatgggtg	accgttctt	cgtgtttac	ggtatcgcc	1140
ctcccatttc	gcagcgcata	gccttctatc	gccttcttga	cgagttttc	ttagttaaa	1200
cagaccacaa	cgggttccct	ctagcggat	caattccgc	cctctccctc	ccccccccc	1260
aacgttactg	gcccgaagccg	cttggaaataa	ggccgggtgt	cgtttgctca	tatgttattt	1320
tccaccat	tgcgttctt	tggcaatgt	agggcccgga	aacctggccc	tgtcttctt	1380
acgagcatc	ctagggtct	ttcccccttc	gccaaaggaa	tgcaaggct	gttgaatgtc	1440
gtgaaggaag	caggcttctt	ggaagcttct	tgaagacaaa	caacgtctgt	agcaccctt	1500
tgcaggcagc	ggaacccccc	acctggcagc	aggtgcctt	gcggccaaaa	gccacgtgt	1560
taagatacac	ctgaaaggc	ggcacaaccc	cagtgcac	ttgtgagtt	gatagttg	1620
gaaagagtca	aatggcttc	ctcaagcgta	ttcaacaagg	ggctgaagga	tgcccagaag	1680
gtacccatt	gtatggatc	tgatctgggg	cctcgggtc	catgctttac	atgtgtttag	1740
tcgaggttaa	aaaacgtcta	ggccccccga	accacgggga	cgtgttttc	cttgaaaaa	1800
cacgataata	ccatggcgc	tattacggcc	tactccaa	agacgcgagg	cctacttggc	1860
tgcatcatca	ctagcctac	aggccggac	aggaaccagg	tcgaggggga	ggtccaagt	1920
gtctccaccc	caacacaatc	tttcttgc	acctgcgtca	atggcgttgc	tttgactgtc	1980
tatcatgttg	ccgcgtcaa	gacccttgc	ggcccaaagg	gcccatacac	ccaaatgtac	2040
accaatgtgg	accaggaccc	ctgcggctgg	caaqcgcccc	ccggggcgcg	ttcttgaca	2100
ccatgcaccc	gcggcagctc	ggaccttac	ttggtcacga	ggcatgccc	tgtcatccg	2160
gtgcggccgc	ggggcagacag	cagggggagc	ctactctcc	ccaggcccg	ctctacttg	2220
aagggcttt	cgggcgttcc	actgctctgc	ccctcggggc	acgctgtgg	cattttcgg	2280
gctgccgtgt	gcacccgagg	ggttgcgaag	gcccgtggact	ttgtaccctg	cgagtctatg	2340
gaaaccacta	tgcgttcccc	ggtttcaacg	gacaactctg	ccctccccc	cgtaccgcag	2400
acatccagg	tggcccaatct	acacggccct	actggtagcg	gcaagagcac	taaggtgc	2460
gctcgatgt	caggccaagg	gtataagggt	tttgtcttgc	acccgtccgt	cgcccccacc	2520
ctaggtttcg	gggcgtatata	gtctaaggca	cattgtatcg	accctaata	cagaatcg	2580
gtaaggacca	tcaccacggg	tgccccatc	acgtactcca	cctatggca	tttcttgcc	2640
gacgggtgtt	gctctggggg	cgcctatgac	atcataat	gtgatgagtg	ccactcaact	2700
gactcgacca	ctatcctggg	catcgccaca	gtcctggacc	aagcggagac	ggctggagcg	2760
cgactcgctg	tgctcgccac	cgctacgc	ccgggatcg	tcaccgtgc	acatccaaac	2820
atcgaggagg	ttgtctgtc	cagcactga	gaaatcccc	tttatggca	agccatcccc	2880
atcgagacca	tcaaggggg	gaggcaccc	atttctgc	atccaaagaa	gaaatgtat	2940
gagctcgcc	cgaactgtc	cggcctcgga	ctcaatgc	tagcatatta	ccggggcctt	3000
gatgtatccg	tcatccaac	tagcggagac	gtcattgtcg	tagcaacgg	cgctctaatg	3060
acgggctta	ccgggtactt	cgactcag	atcgactgc	atacatgtgt	caccagaca	3120
gtcgacttca	gcctggaccc	gacccatc	attgagacga	cgaccgtgc	acaagacgc	3180
gtgtcacgt	cgcacggcg	aggcaggact	ggtagggca	ggatgggc	catttacagg	3240
gtgactccag	gagaacggcc	ctcgccat	ttcgattct	cggttctgt	cgagtgtat	3300
gacgcgggct	gtgttggta	cgagctac	ccccccgaga	cctcagttag	tttgccgg	3360
tacctaaca	caccagggtt	gcccgttgc	cagaccatc	tggagtttgc	ggagacgc	3420
tttacagggc	tcacccacat	agacgccc	ttcttgc	agactaa	ggcaggagac	3480
aactccct	acctggtagc	ataccagct	acgggtgc	ccagggtc	ggctccaccc	3540
ccatcggtgg	accaaatgt	gaagtgtctc	atacggctaa	agcctacgt	gcacggccca	3600

acggccctgc tgtataggct gggagccgtt caaaacgagg ttactaccac acacccata 3660  
 accaaataca tcatggcatg catgtcggtc gacctggagg tcgtcacgag cacctgggt 3720  
 ctggtaggcg gagtcctagc agctctggcc gcgtattgcc tgacaacagg cagcgtggc 3780  
 attgtggca ggatcatctt gtccggaaag ccggccatca ttcccgcacag ggaagtccct 3840  
 taccgggagt tcgatgagat ggaagagtgc gcctcacacc tcccttatcat cgaacaggga 3900  
 atgcagctcg ccgaaacaatt caaacagaag gcaatcggtg tgctgcaaac agccaccaag 3960  
 caagcggagg ctgtctgtcc cggtgtggaa tccaagtgcc ggaccatcg agccttctgg 4020  
 gcaagcata tggtaattt caicagcccc atacaatatt tagcaggctt gtccactctg 4080  
 cctggcaacc cccgatagc atactgtatgc gattcacag cctctatcac cagccgc 4140  
 accacccaaac ataccctctt gtttaacatc ctggggggat gggtggccgc ccaacttgct 4200  
 cctcccaagcg ctgttctgc ttctgttaggc gccggcatcg ctggagcggc tggggcagc 4260  
 ataggccttgc ggaagggtgt tggtggatatt ttggcagggt atggagcagg ggtggcaggc 4320  
 gcgctcggtt ccttaaggt catgagcggc gagatgcctt ccacccgagga cctggtaac 4380  
 ctactccctgc ctatcccttc ccctggccgc ctatgtcg gggtcgtgtg cgcagcgata 4440  
 ctgcgtccgc acgtggggcc ttcgcgttgc gggtaacca gcacgtgtca ctcaagatctt  
 cagtgatca acgaggactg gattggatat gcaacgggtt cgatctccccc accatca  
 gactgatttca aagacctggc tctctagtttcc accatca tcgtgctgaa gaggtttgg 4680  
 gactgatttca aagacctggc tccagtttca gctctgtcc 4740  
 ctccacgcca tgctccggct cgtggctaag agatgtttgg 4800  
 ctgtttcatgtt caacgtgggt acaaggaggat ctggggggc 4860  
 ctgcccattgt gggcacaga tcacccgaca tggtaaaaaac 4920  
 gcttaggacc tgtagtaaca cgtggcatgg aacattcccc 4980  
 attaacgcgtt acaccacggg cccctgcacg ccctcccccgg  
 ctgtggccgg tggctgtga gggatgtgg gggatggggat ttttccactac 5040  
 gtgacggca tgaccactga ttcacagaag tggatgggg 5100  
 cgggttgcac aggtacgctc ggcgttgcac acccctctta 5160  
 cgggaggagg tcacattctt ggtcgggctc aatcaatacc 5220  
 tgcgagcccg aaccggatgt agcagtgttc acttccatgc  
 acggcggaga cggctaagcg taggctggcc agggatctc 5280  
 tcagcttagcc agtgtgtcgcc gccttcctt aagggcaacat 5340  
 cggacgtcg acctcatcg gccaacactt ctgtggccgc 5400  
 accccgtgg agtcagaaaa taaggtagta attttggact  
 gaggaggatg agagggaaatg cttcgatcc ggcggatgtc 5460  
 cctcgagcga tgcccatatg ggcacccccg gattacaacc 5520  
 aaggaccagg actacgtccc tccatgttgc caccatcc  
 cctccgatac caccatcgac gggatggggatc tggggatggc 5580  
 tctgccttgc cggagctcgcc agccggacgg cttcccgacg 5640  
 gtcggatgtt actctccatg gccccccctt gggatccgac 5700  
 gacgggttgc ttgttaccgt aagcgaggag gctatgtgagg  
 tcctacacat ggacaggcgc cctgtatca ccatgcgtg 5760  
 atcaatgcac tgagcaactc tttgtccgtt caccacaact  
 cgcagcgaa gcctgcggca gaagaagggtc acettgtaca 5820  
 cactaccggg acgtgtctaa ggagatgaag gcaaggcg 5880  
 ctatccgtgg aggaagcctg taagctgacg ccccccacatt  
 tatggggcaa aggacgtccg gaaacctatcc agcaaggccg  
 tggaaaggact tgctggaga gggatggggatc cggggatgtt  
 ccagattinggg gggatcggtt gtgcgagaaa atggccctt  
 cctcaggccg tgatggctc ttcatcgatc ttccaaatact  
 ttctgggtga atgcctggaa agcgaagaaa tgccctatgg  
 tggggacttgc caacggtca cactggatgac atccgtgtt  
 tggacttgc ccccccgaagc cagacaggcc ataaggctgc  
 gggggccccc tgactaattc taaaggccgag aactgcggct  
 ggtgtactga cggacgtcg cggtaatacc ctacatgtt  
 tgcgagctg cgaagctcca ggactgcacg atgtctgtat  
 atctgtgaaa ggcggggac ccaagaggac gaggcgagcc  
 atgacttagat actctggccc ccctggggac ccggccaaac  
 ataacatcat gtcctccaa tggatcgatc ggcacatgc  
 tatctcaccc gtgacccac caccatccatc ggcgggctg  
 actccagtc attcctggct aggcaacatc atcatgtatg  
 atgatccatc tgactctt cttctccatc cttctgtatc  
 ctatgtgtc agatctacgg ggcctgttac tccatggac  
 attcaacgac tccatggct tagcgtatc tcactccatc  
 aatagggtgg ctcatgtcc caggaaactt ggggtaccgc  
 cggggccagaa gtgtccgc taggctactg tcccaagggg  
 aagtaccttctc tcaactgggc agtaaggacc aagtcata  
 aacatccatc cccggctg 7620

CA 02303526 2000-03-31

tcccagttgg atttatccag ctgggtcggt gctgggttaca gggggggaga catatatcac 7680  
agcctgtctc gtggccgacc ccgcgtggtc atgtggtgcc tactcctact ttctgttaggg 7740  
gtaggcatct atctactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800  
ccatcctgtt ttttccctt ttttttttc tttttttttt tttttttttt tttttttttt 7860  
ttttctcctt ttttttccctt cttttttcc tttttttcc tttggtgtggct ccatcttagc 7920  
ccttagtcacg gctagctgtg aaagggtccgt gagccgcgttg actgcagaga gtgctgatac 7980  
tggcctctct gcagatcaag t 8001

<210> 9  
<211> 11076  
<212> DNA  
<213> Hepatitis C Virus

<400> 9	ggcagcccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttatgtatgag tgtcgtgcag cctccaggac 120	
ccccccctccc gggagagcca tagtggctcg cgaaaccggg gagtacaccg gaattgccag 180	
gacgaccggg tcctttcttg gcgagactgc tagccgaga gtgttgggtc cgaaaaggcc ggagatttg gcgtgcccc 240	
gtgcttgcga gtgccccggg ctcaaaagaa aaccaaactc aggtctcgta gaccgtgcac ttgtgttact gcctgtatagg 300	
ctcaaaagaa aaccaaactc acgcagggtc tccggccgct gatcaaccgg ctcataatgcct catgagacagc aatcttaaac 360	
caatcggtc ctctgtatgcc ttgtcaagac cgacctgtcc cgtggcttgc cacgacgggc gaccaaccg ggcgcgcccc gattgaacaa gatggattgc 420	
gaagggactg gctgttattc ctccgtccgaa gaaagtatcc cggctactcg cccattcgac tgggtggaga ggctatttogg ctatgactgg gcacaacaga 480	
cggctactcg tcttgcgtat ccgaactctg cgcagggtc aacaccaacg gggggcggcat tgggtggaga ggctgtcage gcaggggcgc ccggttctt 540	
ttgaagccgg tcttgcgtat ccgaactctg cgcagggtc atgaacttgc ggtccctgaa atgaacttgc ggtccctgaa atgaacttgc 600	
ccgaactctg cgcagggtc aatatcatgg tgggtggaga ggctgtcage gaccaagggc aacatcgcat gaccaacaga 660	
atggcgatgc ctgttgcggc actgtggccg gctgggtgtt ttgtcaaga gcttggccgc ctcccgatttc gcaagcgcattt cgcaggatctc actgaagcgg 720	
actgtggccg gctgggtgtt ttgtcaaga gcttggccgc ctcccgatttc gcaagcgcattt cgcaggatctc actgaagcgg 780	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 840	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 900	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 960	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 1020	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 1080	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 1140	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 1200	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 1260	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 1320	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 1380	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 1440	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 1500	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 1560	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 1620	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 1680	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 1740	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 1800	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 1860	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 1920	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 1980	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 2040	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 2100	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 2160	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 2220	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 2280	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 2340	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 2400	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 2460	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 2520	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 2580	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 2640	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 2700	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 2760	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 2820	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 2880	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 2940	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 3000	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 3060	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 3120	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 3180	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 3240	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 3300	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 3360	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 3420	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 3480	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 3540	
ccggctactcg tccaccatatt tgccgttcc acggcatttc cttaggggtt gcaatccggc ctttggataaa ggccgggtgtt gacccgttgc accccgtata 3600	

tgggttcgg ggccctgggtt gacaccaga tgcttggcc actaccata caggcttgg 3660  
 cactaccct gcactgtcaa ctttaccatc ttcaaggta ggatgtacgt ggggggagtg 3720  
 gagcacaggc tcgaagccgc atgcaattgg actcgaggag agcgtttaa cctggaggac 3780  
 agggacagat cagagcttag cccgctgtg ctgtctacaa cgagtgca ggtattgcc 3840  
 tggatgtcc caaccctacc ggctctgtcc actgggttga tccatctcca tcagaacgtc 3900  
 gtggacgtac aatacctgtt cggataggg tccgggttg tctccttgc aatcaaattgg 3960  
 gagtatgtcc tggatgtcc ctttctctg gggacgcgc gctgtgtc ctgcttgg 4020  
 atgatgtgc tggatgtca agctgaggcc gcccstagaga acctgggttgc cctcaacgcg 4080  
 gcatccgtgg ccggggcgca tggatttcc tccctctcg tggatgtcc tggatgtcc 4140  
 tacatcaagg gcaggctggc ccctggggcg gcatatgccc tctacggcgt atggccgta 4200  
 ctccgtcc tggatgtcc accaccacga gcatacgcca tggaccggga gatggcagca 4260  
 tcgtgcggag ggcgggttt ctaggtctg atacicttg a cttgtcacc gcaactataag 4320  
 ctgttcctcg tggatgtcc atggatgttca caatattttt tcaccaggcgc cgaggcacac 4380  
 tggatgtgt ggatcccccc cctcaacgtt cggggggggcc gcgatggcgt catcctctc 4440  
 acgtgcgcgca tccaccaga ggttccactca tggatgtcc ggcgggtata accaaatggc 4500  
 gggctcatc gtgcgtcat gctctcatga agtggccgc actgacaggt acgtacgtt tggatgtcc tggatgtcc 4560  
 cgggacttggg cccacgcggg tctacgagac cttgggttgc ttttgcgttcc 4620  
 tctgatatgg agaccaaggat atcttggcc tggatgtcc ggcgggttgc aatccatcc 4680  
 atggatgtcc tggatgtcc gggatgttgc gctgtgtcc aagttgtctg gggatgttgc caccatcttcc 4740  
 cggggggcc tggatgtcc gggatgttgc tggatgtcc tggatgtcc 4800  
 atggatgtcc tggatgtcc gggatgttgc tggatgtcc tggatgtcc 4860  
 agccttgcggcc tggatgtcc gggatgttgc tggatgtcc tggatgtcc 4920  
 cgaggccatc tggatgtcc gggatgttgc tggatgtcc tggatgtcc 4980  
 gggggggcc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 5040  
 gggggggcc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 5100  
 atcaccctaa tggatgtcc tggatgtcc tggatgtcc tggatgtcc 5160  
 ggcgttcc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 5220  
 gccgatgtca tccgggtcc tggatgtcc tggatgtcc tggatgtcc 5280  
 cccgtctcttacttgc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 5340  
 gtgggcatct tccgggtcc tggatgtcc tggatgtcc tggatgtcc 5400  
 cccgtcgatct tggatgtcc tggatgtcc tggatgtcc tggatgtcc 5460  
 cccgtcgatct tggatgtcc tggatgtcc tggatgtcc tggatgtcc 5520  
 agcactaagg tggatgtcc tggatgtcc tggatgtcc tggatgtcc 5580  
 tccgtcgcc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 5640  
 aacatcgaaa tggatgtcc tggatgtcc tggatgtcc tggatgtcc 5700  
 ggcaagttt tggatgtcc tggatgtcc tggatgtcc tggatgtcc 5760  
 gagtgccact caactgactc tggatgtcc tggatgtcc tggatgtcc 5820  
 gagcggctg gggcgact tggatgtcc tggatgtcc tggatgtcc 5880  
 gtggccacatc caaacatcgaaa tggatgtcc tggatgtcc tggatgtcc 5940  
 ggcaagttt tggatgtcc tggatgtcc tggatgtcc tggatgtcc 6000  
 tggactcaa tggatgtcc tggatgtcc tggatgtcc tggatgtcc 6060  
 tggactcaa tggatgtcc tggatgtcc tggatgtcc tggatgtcc 6120  
 acggacgttca tggatgtcc tggatgtcc tggatgtcc tggatgtcc 6180  
 tggatgtcc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 6240  
 tggatgtcc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 6300  
 tggatgtcc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 6360  
 tggatgtcc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 6420  
 tggatgtcc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 6480  
 tggatgtcc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 6540  
 aggttgcgtt gggatgttgc tggatgtcc tggatgtcc tggatgtcc 6600  
 gtcgtatcgt gggatgttgc tggatgtcc tggatgtcc tggatgtcc 6660  
 cccgttcaaaa tggatgtcc tggatgtcc tggatgtcc tggatgtcc 6720  
 cccgttcaaaa tggatgtcc tggatgtcc tggatgtcc tggatgtcc 6780  
 tggatgtcc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 6840  
 gaaaggccgc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 6900  
 atggatgtcc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 6960  
 agaaggcaat tggatgtcc tggatgtcc tggatgtcc tggatgtcc 7020  
 tggatgtcc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 7080  
 gcccggataca tggatgtcc tggatgtcc tggatgtcc tggatgtcc 7140  
 tggatgtcc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 7200  
 ccaacatacc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 7260  
 gggatgtcc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 7320  
 tggatgtcc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 7380  
 gcccggataca tggatgtcc tggatgtcc tggatgtcc tggatgtcc 7440  
 tggatgtcc tggatgtcc tggatgtcc tggatgtcc tggatgtcc 7500  
 aaccacgtca tggatgtcc tggatgtcc tggatgtcc tggatgtcc 7560  
 aaccacgtca tggatgtcc tggatgtcc tggatgtcc tggatgtcc 7620

gagagcgacg ctgcagcacg tgtcaactca atccctctca gtcttaccat cactcagctg 7680  
ctgaagggc ttaccagggt gatcaacgag gactgctcca cgccatgctc cggctcgtag 7740  
ctaagagatg tttgggattt gatatgcacg gtgttgaactg atttcaagac ctggctccag 7800  
tccaagctcc tgccgcatt gccgggagtc cccttcttct catgtcaacg tgggtacaag 7860  
ggagtcggc gggcgacgg catcatcaa accacactgcc catgtggggc acagatcacc 7920  
ggacatgtgaaaacgggtt catgaggatc gtggggccta ggacctgttag taacacgtgg 7980  
catggAACAT tccccattaa cgcgtacacc acggggccct gcacgcctc cccggcgtca 8040  
aattttctt gggcgtgtg gggtggct gctgaggagt acgtggaggt tacgggggt 8100  
ggggatttcc actacgtgac gggcatgacc actgacgacg taaagtgcctt gtgcaggtt 8160  
ccggcccccg aattttcac agaagtggat ggggtgggt tgacaggtta cgctccagcg 8220  
tgcaaaacccc tcctacgggaa ggaggtcaca ttcttggctcg ggctcaatca atacctgggt 8280  
gggtcAACAGC tcccatgcga gcccgaaccg gatgttagcag tgctcaactc catgtccacc 8340  
gaccctccc acattacggc ggagacggct aagcgttaggc tggccagggg attcctccc 8400  
cccttggcca gctcatcagc tagccagctg tctggccctg ccttgaaggc aacatgact 8460  
accggctatg actccccgggaa cgtggccatc atcgaggcaca acctctgtg cgcggcaggag 8520  
atgggggggaa acatcaccgg cgtggagatca gaaaataagg tagatattt ggacttcttc 8580  
gagccgtcc aaggcgaggg ggatgagagg gaagtatccg ttccggcggg gatcttgcgg 8640  
agggtccagga aattccctcg agcgatgccc atatgggcac gcccggattt caaccctcca 8700  
ctgttagagt ccttggaaagga cccggactac gtccctccag tggtagacgg gtgtacacgg gtgtccattg 8760  
ccgcctgcca aggccccctcc gataccaccc tcaacggagga agaggacggt tgcgttgc 8820  
gaatctaccg tttttttctgc cttggccggag ctcggccacag agacccctgg cagctccggaa 8880  
tcgtcgcccg tgcacagcgcc cacggcaacg gctcttcctg accagcccts cgacgacggc 8940  
gacgccccggat cgcacgttga gatcccgatc gtcctggctt accgtaaagcg ccttgggggg 9000  
tcgtcgcccg tgcacagcgcc cacatggaca ggcggccctga aggaggctag tgaggacgtc 9060  
gtctgtgtct cgatgttcttca tgcactgagc aactcttttc tcacggccatg cgtcgccgg 9120  
gaaaccaaggc tgcccatcaa tgcactgagc aactcttttc tccgttacca caacttggtc 9180  
tatgttacaa catctcgccag cgcaaggctt cggcagaaga aggttacattt tgacagactg 9240  
cagggtccctgg acgaccacta cccggacgtg ctcaaggaga tgaaggcggaa ggcgtccaca 9300  
gttaaggctt aacttctatc cgtggaggaa gcctgttaagc tgacgcccccc acatttggcc 9360  
agatcttaat ttggctatgg gccaaggac gtcgggaaacc tatccagcaa ggccgtttaac 9420  
cacatccgct cccgttggaa ggacttgcg gaaagacactg agaccaaat tgacaccacc 9480  
atcatggcaa aaaaatgggt tttctgtgtc caaccagaga agggggggccg caagccagct 9540  
cgcccttacgttattttccaga cttgggggtt cgtgtgtgcg agaaaatggc ctttacgt 9600  
gtggcttcca cccctccctca ggccgttgcg ggctcttcat acggatttcca atactctct 9660  
ggacagcgccg tgcagggttctt ggtgaatgcc tggaaaggcga agaaatggcc tatgggcttc 9720  
gcatatgaca cccgctgttt tgactcaacg gtcaactgaga atgacatccg ttttgggggtt 9780  
tcaatctacc aatgttgtga cttggccccc gaagccagac ggcggatcggc cgtcttgcaca 9840  
gagccgtttt acatcgggggg cccctctact aattcttaaag ggcagaactg cggctatcgc 9900  
cggtggcccg cgagcggtgt actgacgacc agctgcggta atacccttac attttacttgc 9960  
aaggccgtcg cggccgtgtcg agtgcgaag ctccaggact gacatgtgt cgtatgoggaa 10020  
gacgacccttgc tgcgttatctg tggaaaggcgg gggacccaaag gggacccggc caaaccagaa 10080  
gccttcacgg aggctatgac tagatactt gccccccctg cggccggcc 10140  
tacgacttgg agttgataaac atcatgttcc tccaaatgtgt cagtcgcgca cgatgcatt 10200  
ggcaaaagggt tgtaactatct caccctgttcc cccaccaccc cccttgcgca ggctgcgtgg 10260  
gagacagacta gacacactcc agtcaattcc tggcttagca acatcatcat gtatgcggcc 10320  
accttggggg caaggatgtat cctgtatgact catttttttccatccat tttttttttt 10380  
caacttggaa aaggccctaga acgactccat ggctttagcg agtactccat tttttttttt 10440  
gacccatctc agatcattca tgggcttca tgcccttagga cccatcttct agtcaggaa 10500  
tctccaggtt agatcaatag cttttttttt tttttttttt 10560  
cgagtcgttgc gacatcgggc cagaagtgtc cgcgttaggc tactgtccca gggggggagg 10620  
gctggcactt gtggcaagta cctttaac tggcagttaa ggaccaagct ctttacgt 10680  
ccaatcccg ctgcgttccca gttggattt tccagctggt tcgttgcgttgg 10740  
ggagacatatac acacacggct gtctcggtcc cgaccccgctt ggttcatgtg gtgcctactc 10800  
ctactttctg taggggttagg catctatcta cttcccaacc gatgaacggg gagactaaaca 10860  
ctccaggccat ataggccatc tttttttttt tttttttttt 10920  
ttttttttttt tttttttttt 10980  
tggccatcttgc ttagcccttag tttttttttt tttttttttt 11040  
agagagtgct gatactggcc tctctgcaga tcaagt tccgtgagcc gtttgcactgc 11076

<210> 10  
<211> 8001  
<212> DNA  
<213> Hepatitis C Virus

<400> 10

```

gccaagccccc gattgggggc gacactccac catagatcac tccccctgtga ggaactactg 60
tcttcacgca gaaagcgtct agccatggcg ttagtatgag tgcgtgcag cctccaggac 120
ccccctcccc gggagagcca tagtggtctg cggAACCGGT gagtacaccg gaattgccag 180
gacgaccggg tcctttcttg gatcaacccg ctaatgcct ggagatttg gcgtgcccc 240
gcgagactgc tagccgaga gtgttgggtc cgaaaaggcc ttgtggact gcctgatagg 300
gtgcttgcga gtggcccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaacc 360
ctcaaagaaa aaccaaacgt aaccaacaac ggcgcgcacat gattgaacaa gatggattgc 420
acgcagggtt tccggccgt tgggtggaga ggttattcg ctatgactgg cacaacaga 480
caatcggtt ctctgatgcc ggggttcc ggctgtcagc cgacggggcgc cogggtctt 540
ttgtcaagac cgacctgtcc ggtgcctgtatgaaactgca ggacgaggca ggcggctat 600
cgtggctggc cacgacgggc gttccttgcg cagctgtgtt cgacgttgtc actgaagcgg 660
gaagggactg gctgttattt ggcgaagtgc cggggcagga tctcctgtca tctcacctt 720
ctccctccga gaaagtatcc atcatggctg atgaaatgcg gcccgtgcac acgcttgcac 780
cggttacccg cccattcgac cacaagega aacatgcacat cgagcgagca cgtaactcgga 840
tggaaaggccg tcttgcgtat caggatgtc tggacgaaaga gcatcagggg ctcgcggccag 900
ccgaaactgtt cgccaggctc aaggcgcga tgccgcacgg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgttgcgt aatatcatgg tggaaatgg ccgcgttttctt ggattcatcg 1020
actgtggccg gctgggtgtg gcgaggccgt atcaggacat aacgttggct acccggtata 1080
ttgctgaaga gcttggccgc gaatgggtg accgttccct cgtgttttac ggtatcgccg 1140
ctcccgattt gcagcgcacat gccttctatc gccttcttgcg ctaggttccctc ccccccctt 1200
cagaccacaa cggttccctt ctgggat caattccgc aacgttactg gccacgtgtaa 1260
aacgttactg gccgaaggccg tccaccat tgcgttctt tggcaatgtg aggccccggg aacctggccc tgcgttttctt 1320
acgacgttccctc ctaggggctt tccccctctc gccaaaggaa tgcaaggctt gttgaatgtc 1380
gtgaaggaaag caggctctt ggaagcttctt tgaagacaaaa acctggccac aggtgcctctt 1440
tgcaggcagc ggaacccccc taagatacac ctgaaaggc ggcacaaccc cagtgccacg aacgttactt 1500
gaaagagtc aatggcttcc ctcaagcgtt ttcacaacagg gtcagggttccat gtcgttcc 1560
gtacccctt gtatggatc tgatctgggg cctcgggtca gtcagggttccat gtcgttcc 1620
tcgaggtta aaaacgtcta ggccccccga accacggggta gtcagggttccat gtcgttcc 1680
cacgataata ccatggcgcc tattacggcc tactccaaac agacgcgagg cctacttggc 1860
tgcatacatca ctagcttccatc aggccgggac aggaacccagg gtcagggttccat gtcgttcc 1920
gtctccaccg caacacaata tttcctggcg acctcggtca atggcggttccat gtcgttcc 1980
tatcatggtg cggctcaaa gacccttgcg ggccaaagg gtcagggttccat gtcgttcc 2040
accaatgtgg accaggaccc ctgcggctgg caagcgcaccc gcccacatcac ccaaatgtac 2100
ccatgcacct gcccgcgtc ggaccttac ttggtcacga ggcacgtccgc tgcattccg 2160
gtgcgcggc gggggcagac cagggggagc ctactctccc ccaggccgt ctcttacttg 2220
aagggtctt cgggggttcc actgtctgc ccctcggggc acgtgttggg catctttccg 2280
gctggccgtgt gcaacccgagg ggttgcgaag gcggtggact ttgtaccctg ctagtctatg 2340
ggaaccacta tgcgggtttcc ggttctcaacg gacaactctg cccctccggc cgtaccggcag 2400
acattccagg tggcccatct acacgccttctt acgttgcgttcc gcaagacac taagggtccg 2460
gctgcgtatg caggccaagg gtataagggtt cttgttctgaa acccggttccgt cgcggccacc 2520
ctaggtttcg gggcgatata gtctaaaggca catggatctg accctaaat cagaatcggg 2580
gtaaggacca tcaccacggg tgccccatc acgtacttccaa cctatggcaa gtttcttgc 2640
gacgggtgtt gctctgggg cgcctatgac atcataatat gtgtatggatg ccactcaact 2700
gactcgacca ctatcttggg catcgccaca gtcctggacc aacgcggagac ggctggagcg 2760
cgactcgctg tgctcgccac cgctacccct ccggatcg tccatgtcc acatccaaac 2820
atcgaggagg tggctctgtc cagcaacttgc gaaatccctt ttatggcaa agccatcccc 2880
atcgagacca tcaagggggg gaggcaccc atttctgcg attccaagaa gaaatgtat 2940
gagctcgccg cgaagctgtc cggccctggaa ctaatgtctg tagcatatta ccggggccctt 3000
gatgtatccg tcataccaaat tagcgaggac gtcatgttgc tagcaacggc cgctctaattg 3060
acgggttta cccggcactt cgactcaatg atcgactgca atacatgtgt caccacagaca 3120
gtcgacttca gcctggaccc gacccttccacc attgagacga cgaccgtgcc acaagacgcg 3180
gtgtcacgct cgcacggcg aggccaggact ggttagggca ggatggcat ttacagggtt 3240
gtgactccag gagaacggcc ctcgggcgtt ttcattccct cgggttctgtg cgactgtat 3300
gacgcgggtt gtgttggta cgactcaatg cccggcaga ctcactgtt gttgcgggtt 3360
tacctaaca caccagggtt gcccgtctgc caggaccatc tggatgttccat gggagacgcg 3420
tttacagcc tcaccacat agacgcccatttcttgcg ttcattgtccat agactaagca ggccaggagac 3480
aacttccctt acctggtagc ataccaggtt acgggtgtcc ccagggttccat ggtccaccc 3540
ccatcggtgg accaaatgtt gaaatgttccat agacgcccatttcttgcg ttcattgtccat agactaagca ggccaggagac 3600

```

acggccctgc tgtataggct gggagccgtt caaaacgagg ttactaccac acaccccccata 3660  
 accaaataca tcatggcatg catgtcggtc gacctggagg tcgtcacgag cacctgggtg 3720  
 ctggtaggcg gagtccttagc agctctggcc gcttattgcc tgacaacagg cagcgtggc 3780  
 atttgtggca ggatcatctt gtccggaaag ccggccatca ttcccgcacag ggaagtccctt 3840  
 taccgggagt tcgatgagat ggaagagtgc gcctcacacc tcccttatcat cgaacaggg 3900  
 atgcagctcg cggaaacattt caaacagaag gcaatcggtg tgcgtcaaac agccaccaag 3960  
 caaaggagg ctgtgtctcc cgtgtggaa tccaaatggc ggaccatcg aegcttctgg 4020  
 gcgaagcata tttgtggaaattt catcagcccc atacaatatt tagcaggctt gtcactctg 4080  
 cctggcaacc cccgatagc atcaactgat gcattcacag cctctatcac cagcccgctc 4140  
 accacccaac ataccctctt gtttaacatc ctggggggat gggtgccgc ccaacttgct 4200  
 cctcccaagcg ctgtttctgs tttcgttaggc gcccgcacatcg ctggagccgc ttttggcagc 4260  
 ataggccctt ggaagggtgt tttgtggatatt ttggcagggtt atggagcagg ggtggcaggc 4320  
 gcgctcggtt cctttaaggtt catgagccgc gagggtccct ccaccggaga ccttggtaac 4380  
 ctactccctg statectctc ccctggccgc ctatgtcg ggttgcgtgt cgacgcata 4440  
 ctgcgtccgc acgtggggcc ttcgcttcgc ggggttaacc ggtatgggg gttgtcgt 4500  
 gacacgatca ctcaagatctt cttctatgtt accatcactt tgcctgagag cgacgctgca 4560  
 cagtgatca acgaggactg ctccacgcca tgctccggct cgtggctaag agatgtttgg 4680  
 gattggatat gacgggtttt gactgattt aagacctggc tccagttcaa gctccctgccc 4740  
 cgattggccgg gagtccctt gacccatca tgccaaaccac ctgcccattt ggacacaga tcaccggaca tttgtaaaaac 4860  
 ggttccatgtt ggatgtgggg gcttaggacc tttgttagtaaca cgttgcattt aacattttccc 4920  
 attaacgcgtt acaccacggg ctgtggccgg taggtgtctga ggagtacgtt gaggttacgc cggccaaatata ttcttagggcg 4980  
 ctgtggccgg taggtgtctga gtgacggca tgaccactga caacgttaaa tgccgtgtc aggttccggc ccccaattt 5100  
 ttcacagaag tggatggggt ggggttgcac aggtacgctc cagcgtgcaa acccttccta 5160  
 cggggaggagg tcacatttctt ggttccggctt aatcaataacc tgggtgggtc acagctccca 5220  
 tgcgagctgtt aaccggatgtt agcagtgtctt acttccatgc tcaccggaccc cttccacattt 5280  
 acggccggaga cggtaaagcg taggtgtcc agggatctt cccccccctt ggcagctca 5340  
 tcagcttagcc agctgtctgc gccttcctt aaggcaacat gcactacccg tcatgactcc 5400  
 cccggacgttgc acctcatgtt ggccaacctt ctgtggccggc aggagatggg cggaaacatc 5460  
 accccgttgg agtcagaaaa taaggtagta attttgactt ctttcgagcc gcttcaagcg 5520  
 gaggaggatg agaggggagt atccgttccg gcccggatcc tgccggaggtc cagaaattt 5580  
 cctcgagcgta tgcccatatg ggcacggccg gattacaacc ctccactgtt aggttccctgg 5640  
 aaggacccgg actacgtccc tccagtgtttaa caccgggtgtc cattggccgc tgccaaggcc 5700  
 cctccgatac caccatcactg gaggaaaggagg acgggtgtcc ttttgcataatc taccgtgtct 5760  
 tctgccttgg cggagctgcg cacagagacc ttcggcactt ccgaatgtc ggcgtcgac 5820  
 acggccacgg caacggcccte teetggacccg ccctccggacg acggccacgg gggatccgac 5880  
 gttgatgtctt acttcctccat gccccccctt gggggggggc cggggatcc cgatctcage 5940  
 gacgggtctt ggttaccgtt aacggggggg gcttagtgagg acgtcgctg ctgtcgatg 6000  
 tcctacacat ggacaggccgc cctgatcactg ccattgcgtt cggaggaaac caagctgccc 6060  
 atcaatgcac tgagcaactt tttgtccgtt caccacaact tggtctatgc tacaacatct 6120  
 cgcagcgcaaa acctggccca gaagaagtc accttigaca gactgcagg tctggacac 6180  
 cactaccggg acgtgtctaa ggagatgaaag gccaaggccgt ccacagttaa ggctaaactt 6240  
 ctatccgtgg aggaaggccctg taagctgtccg ccccccattt cggccagatc taaatttggc 6300  
 tatggggcaaa aggacgttgg gaaacctatcc agcaaggccg ttaaccacat ccgtccgt 6360  
 tggaaaggact tggatggaaa cactggacca ccaatttgaca ccaccatcat ggcaaaaaat 6420  
 gaggttttctt gcttccaaacc agagaagggg gggcccaagc cagctcgctt tattgtattt 6480  
 ccagattttgg ggggttgcgtt gtgcgagaaaa atggccctt acgatgtggt ctccaccctc 6540  
 cctcaggccg tgatggccctt ttccatcggg tttccatactt cttctggaca gcccggcag 6600  
 ttccctggta atgcctggaa agcgaagaaa tggccctatgg gcttgcata tgacacccgc 6660  
 tggggggactt caacgggtcac tgagaatgtt atccgtgttgg aggagtcaat ctaccaatgt 6720  
 tggacttgg ccccccggc acggccatcc cggccatccatc tccacagacg gctttacatc 6780  
 gggggccccc tgactaattt taaaggccgg aactgtcggtt atccggccgtt ccggccgagc 6840  
 ggtgtactgtt cggccatccatc tccacatgtt atttgcggc cgttgcggcc 6900  
 tggccatccatc cggccatccatc tccacatgtt gcttgcata tggccggcc 6960  
 atctgtaaaa ggcgggggac ccaagaggac gaggccgagcc tacggccctt cacggaggct 7020  
 atgacttagat acttcctccat ccttggggac ccggccaaac cagaatacga ttggaggtt 7080  
 ataacatcat gctccctccaa tttgttgcgtt gcccacatgtt catctggcaaa aagggtgtac 7140  
 tatctccatccatc gtcggccatccatc ccccccattt gcccggccgtt cgtggggagac agctagacac 7200  
 actccatccatc atttcctggctt aggcaacatc attcgtgtt ccggccatccatc gtcggccatccatc 7260  
 atgatccatgtt tgacttccatc ctttcgttgcgtt acggccatccatc tggccggcc 7320  
 ctatgttgcgtt agacttacgg ggcctgttac tccatgttgcgtt ctttgcgtt acctcagatc 7380  
 attcaacggac tccatggccctt tagcgcattt tcactccatc gttactctcc aggtggagatc 7440  
 aatagggtgg cttcatggccctt cggggaaactt ggggttccgc ctttgcgtt acgttggagatc 7500  
 cggggccagaa gtttccgcgtt taggtactgtt tcccaagggggg ggaggccgttcc cacttgcgtt 7560  
 aagtacctctt tcaactggc cggggaaactt ggggttccgc ctttgcgtt acgttggagatc 7620

CA 02303526 2000-03-31

tcccagttgg atttatccag ctgggtcggt gctgggttaca gcgggggaga catatatcac 7680  
agectgtctc gtgcccgacc ccgctggttc atgtggtgcc tactcctact ttctgttaggg 7740  
gtaggcatct atctactccc caaccgatga acggggagct aaacactcca ggccaatagg 7800  
ccatcctgtt tttttccctt ttttttttc tttttttttt tttttttttt tttttttttt 7860  
ttttctcctt tttttttccctt ttttttttc tttttttttt tttgggtggct ccatcttagc 7920  
cctagtcaacg gctagctgtg aaagggtccgt gagccgcctt actgcagaga gtgctgatac 7980  
tggcctctct gcaagatcaag t 8001

<210> 11  
<211> 11076  
<212> DNA  
<213> Hepatitis C Virus

<400> 11

```

gccagccccc gattggggc gacactccac catagatcac tccccctgtga ggaactactg 60
tcttcacgca gaaagcgct agccatggcg ttagtatgag tgcgtgcag cctccaggac 120
ccccccccc gggagagcca tagtggtctg cgaaccggg gagtacaccg gaattgccag 180
gacgaccggg tcctttctt gatcaacccg ctaatgcct ggagattgg gcgtgcccc 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtaact gcctgatagg 300
gtgttgcga gtgcggggg aggctctgtc gaccgtgcac catgagacg aatcctaacc 360
c:caaagaaa aaccaaaccgt aaccaaacg ggcgcggcat gattgacaaa gatggattgc 420
acgcaggttc tcggccgcg tgggtggaga ggctattcg ctatgactgg gcacaacaga 480
caatcgctg ctctgatgcc gccgtttcc ggctgtcgc gcagggcgc ccggttctt 540
ttgtcaagac cgacctgtcc ggtgcctgtc atgaactgca ggacgaggca ggcggctat 600
cgtggctggc cacgacgggc gttcctgtcg cagctgtgtc cgacgtgtc actgaagcgg 660
gaagggactg gctgttattg ggcaagtgc cggggcagga ttcctgtca ttcaccc 720
tcctgtccga gaaagtatcc atcatgctg atgcaatgcg gggctgtcat acgcttgatc 780
cggttacgtt cccattcgac caccaagcga aacatcgcat cgagcgacg cgtactcgga 840
tggaaaggcg tcttgcgtat caggatgtc tggacgaaa gcatcgggg ctcgcggccag 900
ccgaactgtt cgccaggctc aaggcgcgc tggccgacgg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgttgcgcg aatatcatgg tggaaaatgg cgcgttttctt ggattcatcg 1020
actgtggccg gctgggtgtg ggggaccgct atcaggacat aacgttggct acccggtata 1080
ttgctgaaga gcttggccgc gaatgggtg accgtttctt cgtgttttac ggtatcgccg 1140
tcctccgattc gcaagcgcattc gccttctatc gccttcttgc ctaggttccctt cccctccctt 1200
cagaccacaa cgggttccctt cttagggatc caattccgcg cctctccctt ccccccctt 1260
aacgttactg gccgaagccg cttggataaa ggccgggtgtg cttttgtctat tatgttattt 1320
tccaccatattt tgccgtctt tggcaatgtg agggccgggaa aacctggccc tttttttttt 1380
acgagcattc cttaggggtct ttccccttc gccaaggaa tgcaaggctt gttaatgtc 1440
gtgaaggaaag cagttcttctt ggaagcttctt tgaagacaaa caacgtctgt aacgaccctt 1500
tgcaggcgcg ggaacccccc acctggcgac aggtgcctct gggccaaaa gccacgtgt 1560
taagatacac ctgcaaaaggc ggcacaaccc cagtgcacg ttgtgagttt gatagttgt 1620
gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
gtacccattt gtatgggtc ttagtctgggg cctcgggtgc catgctttac atgtgttttag 1740
tcgaggtaaa aaaacgtcta ggcccccccg accacggggaa cttgggtttt ctttggaaaa 1800
cacgataata ccatgggcac gaatcttaaa cctcaaaaggaa aaaccaaaacg taacaccaac 1860
cgccgcccac aggacgtcaa gttccggggc gttgggtcaga tgcgtggggc agtttac 1920
ttggccgcga gggcccccag gttgggtgtg cgccgcacta ggaagacttcc cgagcggtcg 1980
caacctcgta gaaggcgaca acctatcccc aaggctcgcc agcccgaggg tagggcttg 2040
gctcagcccg ggttcccttgc gccccttatc ggcaatgagg gtttgggtg ggcaggatgg 2100
tcctgtcac cccgtggctc tcggcctagt tggggccca cggaccccg gcgttaggtcg 2160
cgcaatttgg gtaaggctatc cgataacctc acgtggggct tcggcgatct catgggtac 2220
atcccgctcg tcggccccc cctagggggc gtcggccaggg ccctggcgca tgggtccgg 2280
gttctggagg acggcgtaa ctatgcataa gggaaatgtc cgggttgcctt ctttctatc 2340
ttccttttgtt ctttgcgtc ctgtttgacc atcccgatc cgcgttatga agtgcgcac 2400
gtatccggag ttttgcgtatc cacaacgac tgctccaacg caagcatgtt gtatgaggca 2460
gcggacatga tcatgcatac cccgggtgc gtggccgtcg ttggggagaa caactctcc 2520
cgctgctggg tagcgctcac tcccacgctc gcgccaggaa acgctagctt ccccaactac 2580
acgataacgac gccatgtcgat ttgctcgat gggccggctg ctctctgtc cgctatgtac 2640
gtgggagatc ttcggatc ttttgcgtc tcggcccgac ttgttgcaccc ttcgcctcgc 2700
cgccacgaga cagtagggaa ctgcaatgtc tcaatatac cccggccacgt gacaggtcac 2760
cgatggctt gggatgtatc gatgactgg tcacactac cagccctagt ggtatcgac 2820
ttactccggta tcccaacaagg ttcgtgtatc atgttgggg gggcccatgg gggaggctta 2880
cgccggcccttgc cttactattc catgggtggg aactggctt agttctgtat tttttttttt 2940
ctctttggcg gctttgtacgg gggaaacctt gtgacagggg ggacgtatggc caaaaacacc 3000
ctcggttacgg tttttttttt ttcacccggg tcatccaga aaatccatgt tttttttttt 3060
aacggcaactt ggcacatcaa caggactgcc ctgaaactgtca atgactccctt caacactggg 3120
ttcccttgcgtc cgctgttca cgtgcacaaat ttcacactatc ctggatggcc agagcgatcg 3180
gccagctgca gccccatcgaa cgcgttgcgt cagggtggg gggccatcatc ttacaatgt 3240
tcacacatgtt cggccaggag gccttattgt tggactacg caccggccgc gtgcggatcc 3300
gtacccggcg cgccagggtgtt tggccatgt tttttttttt tttttttttt 3360
gggacgaccgg accgggttcgg cgtccctacg tacagttggg gggagaatgt gacggacgtg 3420
ctgtttttttt acaacacgca ggcggccaa ggcaactgtt ttggctgtac atggatgtat 3480
agcactgggt tcaccaagac gtgcgggggc ccccccgtta acatcggggg gatcggtcaat 3540
aaaacccatgtt cctggcccccac ggactgttcc cgaaaggccac ttacaccaag 3600

```

tgggtttcg ggccttggtt gacaccaga tgcttggtcc actaccata caggctttgg 3660  
 cactaccct gcactgtcaa ctttaccatc ttcaaggta ggatgtacgt ggggggagtg 3720  
 gagcacaggc tcgaagccgc atgcaattgg actcgaggag a诶cgtttaa cctggaggac 3780  
 agggacagat cagagcttag cccgctgctg ctgtctacaa cggagtggca gttattgcc 3840  
 tggcccttca ccaccctacc ggctctgtcc actgggttga tccatctcca tcagaacgtc 3900  
 gtggacgtt aatacctgtt cggtagatgg tcggcggtt tcccttgc aatcaaattg 3960  
 gatgttgc tggatgttcc cttcttgc ggggacgcgc gctgtgtgc ctgttgttgc 4020  
 atatgttgc tgatagctca agctggggc gcccctagaga acctgggtt cttcaacgc 4080  
 gcatccgtgg cggggggcga tggatttctc ttcttctcg tgttttctg tgctgccttgc 4140  
 tacatcaagg gcaaggctggt ccctggggcg gcatatgcc tctacgggt atggccgcta 4200  
 ctccctgtcc tggatgttcc accaccacga gataacgcca tggacggga gatggcagca 4260  
 tcgtgcggag ggcgggttt ctaggtctg atactcttga cttgtcacc gcaactataag 4320  
 ctgttccctcg cttaggtctcat atgggttta caatattttt tcaaccaggc cgagggcacac 4380  
 ttgcaagttt ggtatcccccc cctcaacgtt cggggggggcc gcgatggcg catccttc 4440  
 acgtgcgcga tcaccggaga gtaatctt accatcacca aaatcttgc gcccatactc 4500  
 ggccactacta tggatgttcc gctgttata accaaaatgtc cgtactctg ggcgcacac 4560  
 gggctcattc gtgcattgtat gctgggtcgg aagggtgttgc ggggtcatta tgccttgc 4620  
 gctctcatga agttggccgc actgacaggt actgtacgtt atgaccatct caccctactg 4680  
 cgggacttggg cccacgcggg cctacagagac cttgcgggtt cagttgagcc cgtcgttcc 4740  
 tctgatatgg agaccaagg tttttttt gggggcagaca ccccttgc tggggacats 4800  
 atcttggggcc tggccgttcc cggccgcagg gggagggaga tacatctggg accggcagac 4860  
 agccttgaag ggcagggttgc gcaacttcc ggccttattt cggctactc ccaacagacg 4920  
 cgaggcttcc ttggctgtat ctttttttgc ctttttttgc tggggacatggcc 4980  
 gggggaggtcc aagtggtctc caccgcacca caatcttcc tggccggccca aaaggggccca 5100  
 gtgtgttgg tggatgttcc tggatgttcc tggatgttcc tggccggccca aaaggggccca 5160  
 ataccaccaaa ttttttttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 5220  
 ggcgttccct ttttttttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 5280  
 gccgatgtca ttttttttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 5340  
 cccgttccct acttggatgttcc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 5400  
 gtggggatct ttttttttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 5460  
 cccgtcggat ttttttttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 5520  
 cccgttccct ttttttttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 5580  
 agcaactaagg ttttttttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 5640  
 tccgtcggcc ccccttccat ttttttttgc ttttttttgc ttttttttgc ttttttttgc 5700  
 aacatcaaaaa ttttttttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 5760  
 ggcaagtttcc ttttttttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 5820  
 gagtgccact caactgtact gaccatata ctggcatcg gcacagtctt ggaccaagcg 5880  
 gagacgcgtg gaggcgcact ctttttttgc ttttttttgc ttttttttgc ttttttttgc 5940  
 gtgcacatcc caaacatcgat ttttttttgc ttttttttgc ttttttttgc ttttttttgc 6000  
 ggcaagccca tcccatcgat gaccatcaag gggggggggc accttattt ctgcatttc 6060  
 aagaagaaat gtatgtatgt ttttttttgc ttttttttgc ttttttttgc ttttttttgc 6120  
 tattaccggg gcttggatgttcc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 6180  
 acggacgttcc taatgacggg ttttttttgc ttttttttgc ttttttttgc ttttttttgc 6240  
 ttttttttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 6300  
 gtgcccacaag acggcgttcc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 6360  
 ggcatttaca ggtttgttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 6420  
 ctgtcgagat gctatgttcc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 6480  
 gtttttttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 6540  
 ttcttgggaga gcttggatgttcc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 6600  
 aagcaggcag gagacaactt ccccttccat ttttttttgc ttttttttgc ttttttttgc 6660  
 gtcaggctc caccctccat ttttttttgc ttttttttgc ttttttttgc ttttttttgc 6720  
 acgctgcacg gggccaaacgc accacacacc ccataaccaa atacatcatg gcatgtatgttcc ttttttttgc 6780  
 acgaggacact ggggtgttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 6840  
 acaggcagcg ttttttttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 6900  
 gacagggaaag ttttttttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 6960  
 tacatcgaccc ttttttttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 7020  
 caaacagcca ccaagcaagg gggaggctgttcc ttttttttgc ttttttttgc ttttttttgc 7080  
 atcgaaagctt ttttttttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 7140  
 ggcttggatgttcc ttttttttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 7200  
 atcaccaccc cggccatccat ccccttccat ttttttttgc ttttttttgc ttttttttgc 7260  
 gggcccaacaa ttgttcccttcc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 7320  
 ggggggttgc gggccatccat ttttttttgc ttttttttgc ttttttttgc ttttttttgc 7380  
 gcaagggttgg cgggggggttcc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 7440  
 gaggacacttgg ttttttttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 7500  
 gtgtgcgcag cggccatccat ttttttttgc ttttttttgc ttttttttgc ttttttttgc 7560  
 aaccggctga taggttgc ttttttttgc ttttttttgc ttttttttgc ttttttttgc 7620



## CLAIMS

1. A hepatitis C virus (HCV) cell culture system, which comprises mainly eukaryotic cells containing transfected HCV specific genetic material, characterized in that,  
5 the eukaryotic cells are human hepatoma cells and the transfected HCV specific genetic material is a HCV-RNA construct, which comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 10 3' NTR as well as an additional marker gene for selection (selection gene).
2. A cell culture system according to Claim 1, 15 characterized in that,  
the hepatoma cells are derived from a usual hepatoma cell line.
3. A cell culture system according to Claim 1, 20 characterized in that,  
the hepatoma cells are extracted from a primary hepatoma cell culture.
4. A HCV-RNA construct, characterized in that,  
25 it comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR as well as an additional marker gene for selection (selection gene).
5. A HCV-RNA construct according to Claim 4, 30 characterized in that,  
it comprises one of the nucleotide sequences illustrated in the sequence protocols SEQ ID NO: 1 to SEQ ID NO: 11.
- 35 6. A HCV-RNA construct according to Claim 4, characterized in that,

the 3' NTR has a nucleotide sequence, which has been selected from the group of nucleotide sequences (a) to (i) listed in the following:

- 5           (a) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTA  
              GCTTTTTTTTTCTTTTTTGAGAGAGAGAGTCTCACTCTGTTGCC  
              CAGACTGGAGT
- 10          (b) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTA  
              GTCTTTTTTTCTTTTTTGAGAGAGAGAGTCTCACTCTGTTGCCA  
              GACTGGAGC
- 15          (c) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTAAAT  
              CTTTTTTTTTCTTTTTTGAGAGAGAGAGTCTCACTCTGTTGCCA  
              GACTGCAGC
- 20          (d) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTT  
              AGTCTTTTTTTCTTTTTTGAGAGAGAGAGTCTCACTCTGTTGC  
              CCAGACTGGAGT
- 25          (e) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTA  
              GTCTTTTTTTCTTTTGAGAGAGAGAGTCTCACTCTGTTGC  
              CCAGACTGGAGT
- 30          (f) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTA  
              GTCTTTTTTTCTTTTGAGAGAGAGAGTCTCACTCTGTTG  
              CCCAGACTGGAGT
- 35          (g) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTA  
              GTCTTTTTTTCTTTTGAGAGAGAGAGTCTCACTCTGTTGCC  
              CAGACTGGAGT
- (h) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTT  
              TAATCTTTTTTTCTTTTGAGAGAGAGAGTCTCACTCTGT  
              TGCCCAGACTGGAGT

(i) ACGGGGAGCTAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTTT  
AATCTTTTTTTTTTTCTTTTTTGAGAGAGAGAGTCTCACTCTG  
TTGCCAGACTGGAGT

- 5 7. A HCV-RNA construct according to one of the Claims 4 to 6, characterized in that,  
the marker gene for selection is a resistance gene and preferably an antibiotic resistance gene.
- 10 8. A HCV-RNA construct according to one of the Claims 4 to 6, characterized in that,  
the marker gene for selection is a neomycin phosphotransferase gene.
- 15 9. A HCV-RNA construct according to one of the Claims 4 to 8, characterized in that,  
the marker gene for selection is integrated in the HCV-RNA downstream from the 5' NTR.
- 20 10. A HCV-RNA construct according to one of the Claims 4 to 83, characterized in that,  
the marker gene for selection is linked to the HCV-RNA via a ribozyme or a recognition sequence for a ribozyme.
- 25 11. A HCV-RNA construct according to one of the Claims 4 to 10, characterized in that,  
it has an integrated reporter gene.
- 30 12. A HCV-RNA construct according to Claim 11, characterized in that,  
the reporter gene is a gene from the group of the luciferase genes, the CAT gene (chloramphenicol acetyl transferase gene), the lacZ gene (beta galactosidase gene), the GFP gene (green fluorescence protein gene), the GUS gene (glucuronidase gene) and the SEAP gene (secreted alkaline phosphatase gene).

13. A HCV-RNA construct according to one of the Claims 4 to 11, characterized in that,  
their replication has an influence on the expression  
of a (cellular) surrogate marker gene.  
5
14. A HCV-RNA construct according to one of the Claims 11 to 13, characterized in that,  
the resistance gene is cloned into the open reading frame of the HCV-RNA in such a way that it will only be transferred to an active form after proteolytic processing.  
10
15. A HCV-RNA construct according to one of the Claims 11 to 14, characterized in that,  
the reporter gene and the marker gene for selection are arranged in the construct in such a way, that they are expressed as a fusion protein.  
15
16. A cell culture system according to one of the Claims 1 to 3, characterized in that,  
the HCV-RNA construct is a construct according to at least one of the Claims 4 to 15.  
20
17. A cell culture system according to Claim 1,  
characterized in that,  
the cells containing the HCV-RNA construct are lodged at the DSMZ, Braunschweig, Germany under the deposit number DSM ACC2394 (laboratory name HuB1 9-13).  
25
18. A use of a cell culture system according to one of the Claims 1 to 3 or 16 to 17 and/or a HCV-RNA construct according to one of the Claims 4 to 15 for the production and/or evaluation and/or testing of treatments and/or diagnostics for the treatment of HCV infections in particular.  
30  
35

19. A use of a cell culture system according to one of the Claims 1 to 3 or 16 to 17 and/or a HCV-RNA construct according to one of the Claims 4 to 15 for the production of a vaccine against HCV infections.

5

20. A use a HCV-RNA construct according to one of the Claims 4 to 15 for the production of a liver cell specific vector in gene therapy.

10 21. A HCV-RNA construct according to one of the Claims 4 to 15; characterized in that, it has an integrated foreign gene and can be used to transfect this foreign gene in a target cell, which can be used to express this foreign gene.

15

22. A process for the production of cell culture adapted mutants of a HCV-RNA construct according to one of the claims 4 to 15, in which the mutants have increased replication efficiency compared to the HCV-RNA construct,

20

characterized in that, a cell culture system according to claim 1; in which the transfected HCV specific genetic material is a HCV-RNA construct with a selection gene according to one of the claims 4 to 15, is cultivated on/in the selection medium corresponding to the selection gene, that the cell clones created are collected, and that the HCV-RNA constructs or parts thereof are isolated from these cell clones.

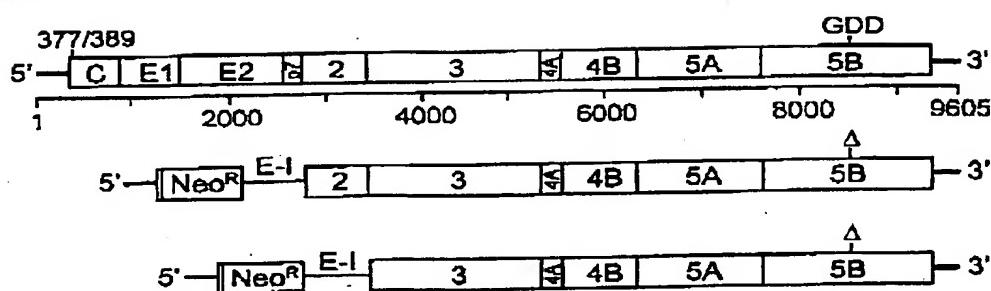
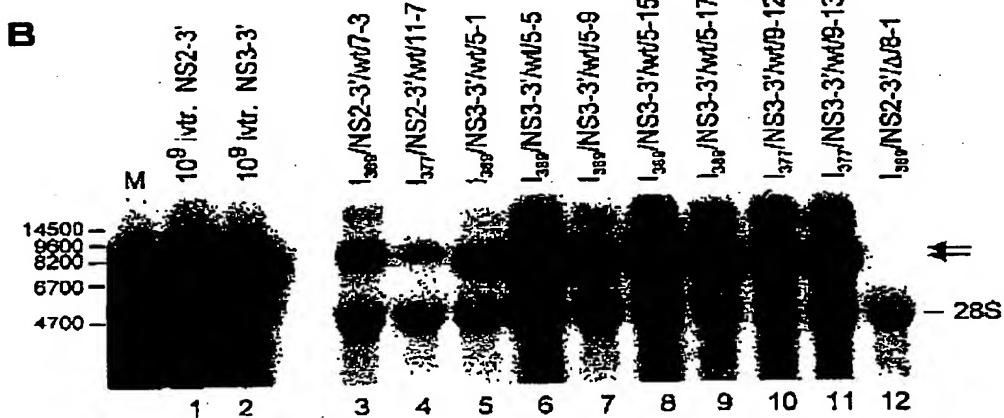
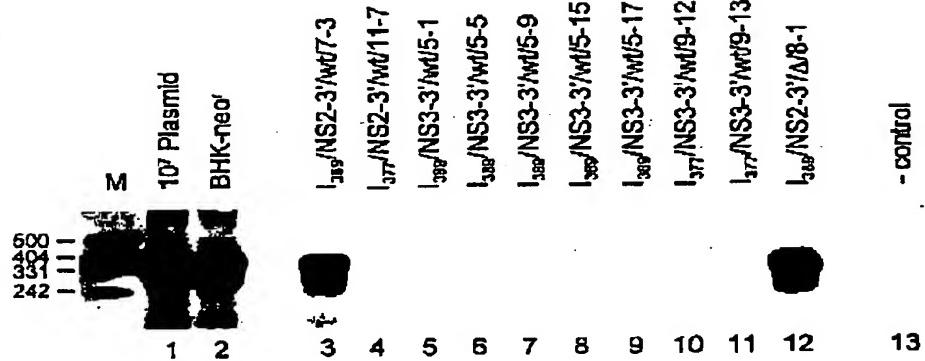
25

30 23. A process according to claim 22, characterized in that, the isolated HCV-RNA constructs are passaged at least one more time. That is, they are transfected in cells of a cell culture system according to claim 1. The cell culture system of claim 1 thereby obtained, in which the transfected HCV specific genetic material is

35

- the isolated HCV-RNA construct with a selection gene,  
is cultivated on/in the selection medium corresponding  
to the selection gene, the cell clones created are  
collected, and the HCV-RNA constructs are isolated  
from these cell clones.
- 5
24. A process for the production of mutants of a HCV full-length genome or of a HCV subgenome or any HCV construct with increased replication efficiency in  
10 comparison to the original HCV full-length genome or subgenome or HCV-RNA construct,  
characterized in that,  
a cell culture adapted mutant of a HCV-RNA construct  
is produced and isolated by a process according to  
15 claim 22 or 23,  
the nucleotide and amino acid sequence of these  
mutants is determined and the type, number and  
positions of the nucleotide and amino acid mutations  
is determined by comparison with the nucleotide and  
20 amino acid sequence of the original HCV-RNA construct,  
and these mutations are introduced in an (isolated)  
HCV full-length genome or a HCV subgenome or any HCV-  
RNA construct, either by site-directed mutagenesis or  
by exchange of fragments containing the relevant  
25 mutations
25. A cell culture adapted HCV-RNA construct with high  
replication efficiency,  
characterized in that,  
30 it is derivable from a HCV-RNA construct according to  
one of the claims 4 to 15 through nucleotide and/or  
amino acid mutations and that it is obtainable with a  
process according to one of the claims 22 to 24.
- 35 26. A cell culture adapted HCV-RNA construct according to  
claim 25,  
characterized in that,

- 5 it carries one or several of the following amino acid exchanges: 1283 arg -> gly and/or 1383 glu -> ala and/or 1577 lys -> arg and/or 1609 lys -> glu and/or 1936 pro -> ser and/or 2163 glu -> gly and/or 2330 lys -> glu and/or 2442 ile -> val.
- 10 27. A cell culture adapted HCV-RNA construct according to claim 25 or 26,  
characterized in that,  
it displays one or several of the nucleotide and/or amino acid exchanges shown in Table 3, where Table 3 is an integral part of this claim.
- 15 28. Cell culture adapted mutant of a HCV-RNA construct or of a HCV full-length genome with increased replication efficiency in comparison to the original HCV-RNA construct or original HCV full-length genome,  
characterized in that,  
it is obtainable with a process in which the number and type of mutations in a cell culture adapted HCV-RNA construct according to claim 24 are determined through sequence analysis and sequence comparison, and these mutations are introduced in a HCV-RNA construct, particularly in a HCV-RNA construct according to one  
25 of the claims 4 to 15, or in an (isolated) HCV-RNA full-length genome, either by targeted mutagenesis or by exchange of sequence sections containing the relevant mutations.
- 30 29. Hepatitis C virus particles or virus-like particles,  
characterized in that,  
they are obtainable by a process according to one of the claims 22 - 24.
- 35 30. Cells infected with Hepatitis C virus particles or virus-like particles, according to claim 29.

**A****B****C****Fig. 1**

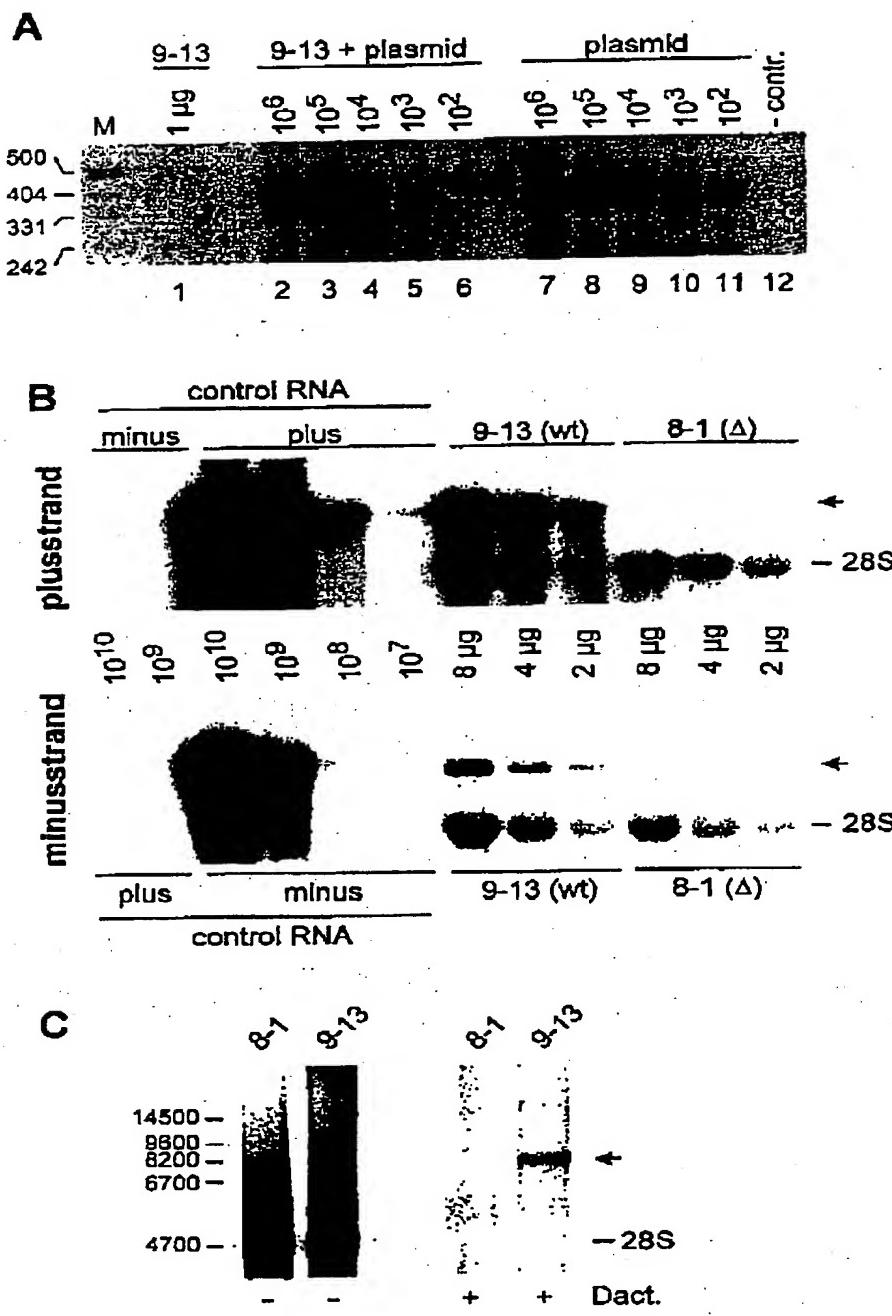
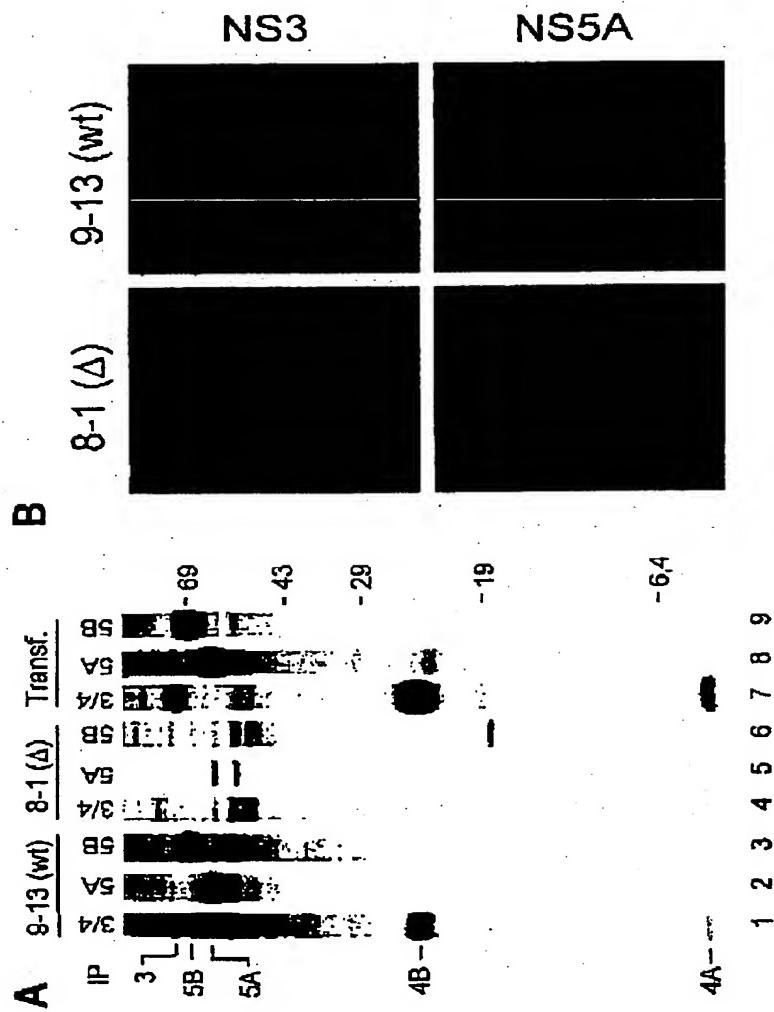


Fig. 2

Fig. 3



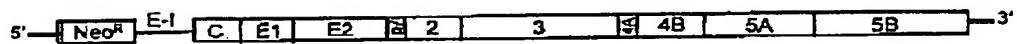
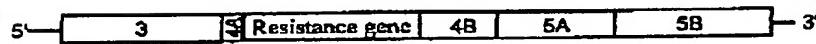


Fig. 4

A



B

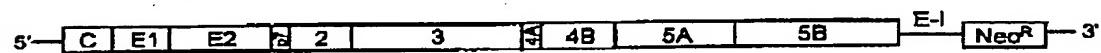
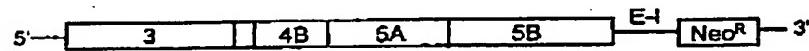
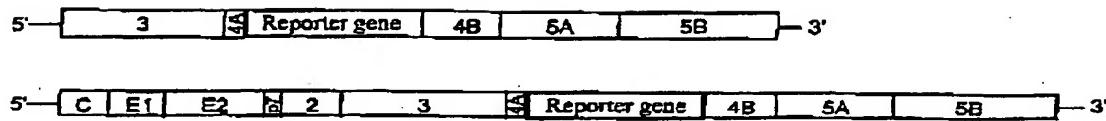
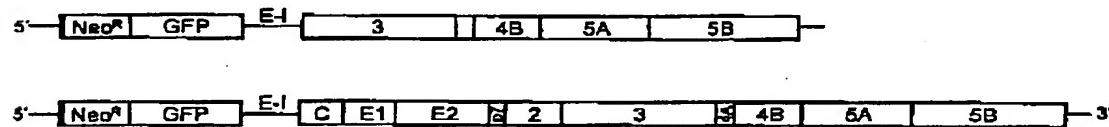


Fig. 5

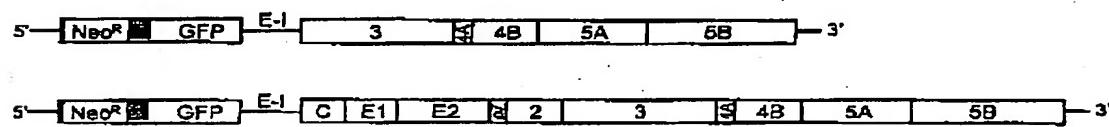
**A**



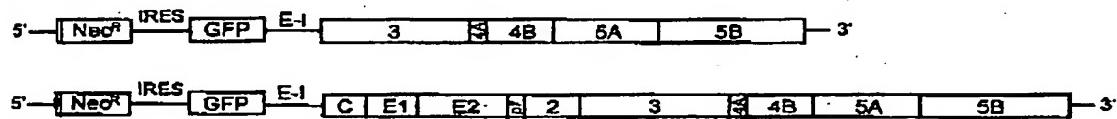
**B**



**C**



**D**



**Fig. 6**

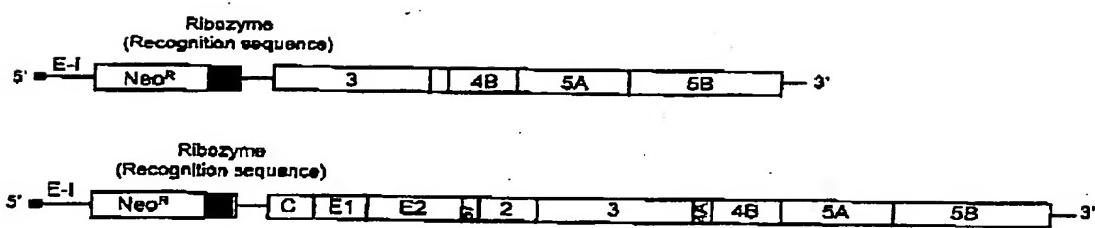


Fig. 7

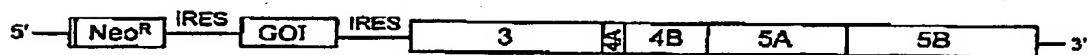


Fig. 8

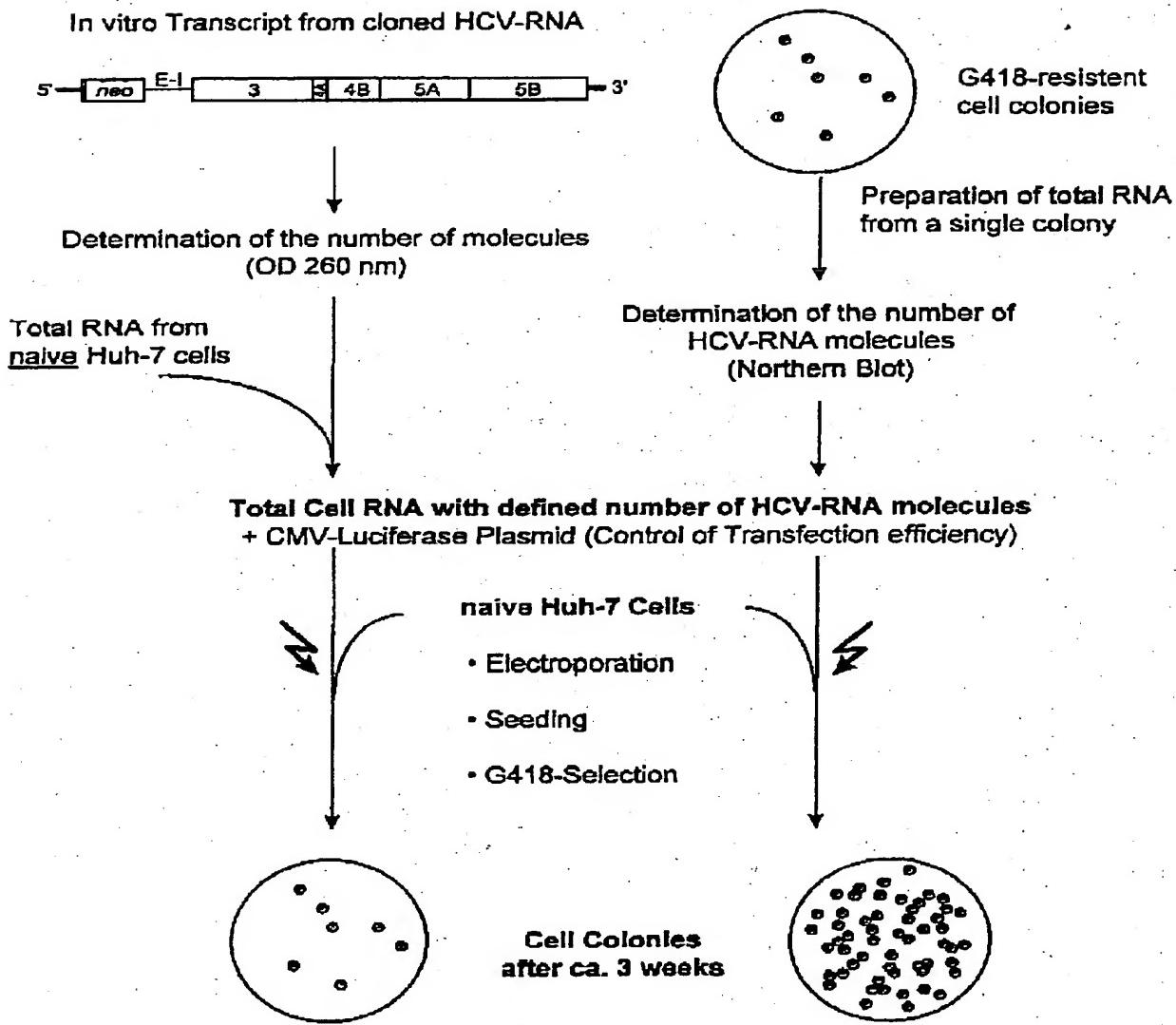


Fig. 9

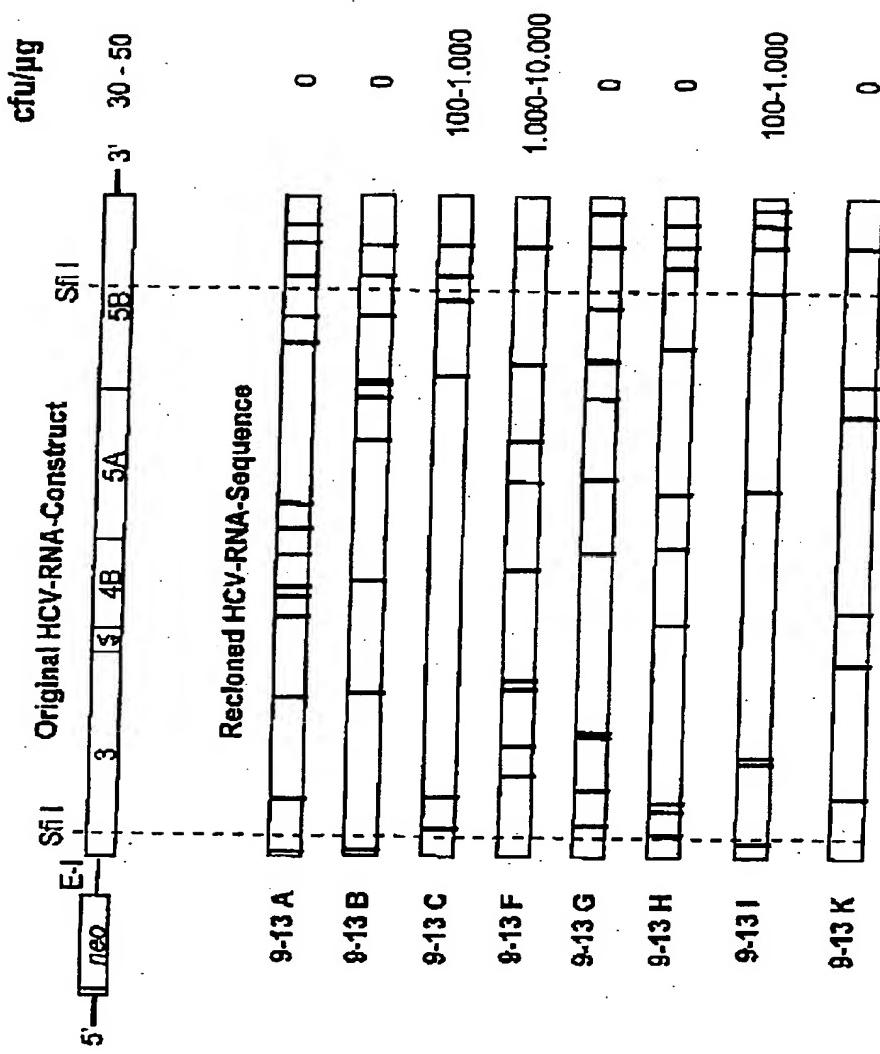
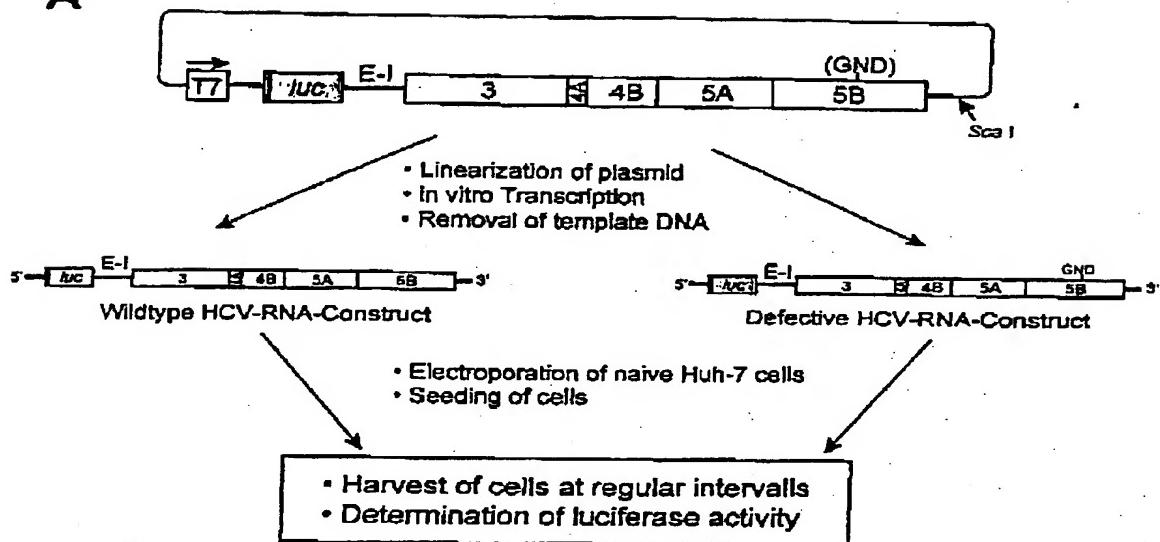
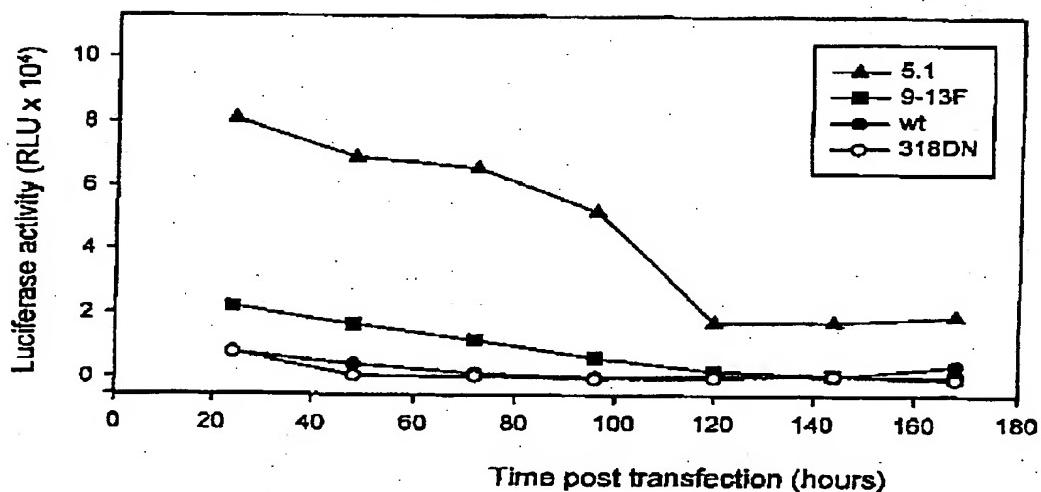
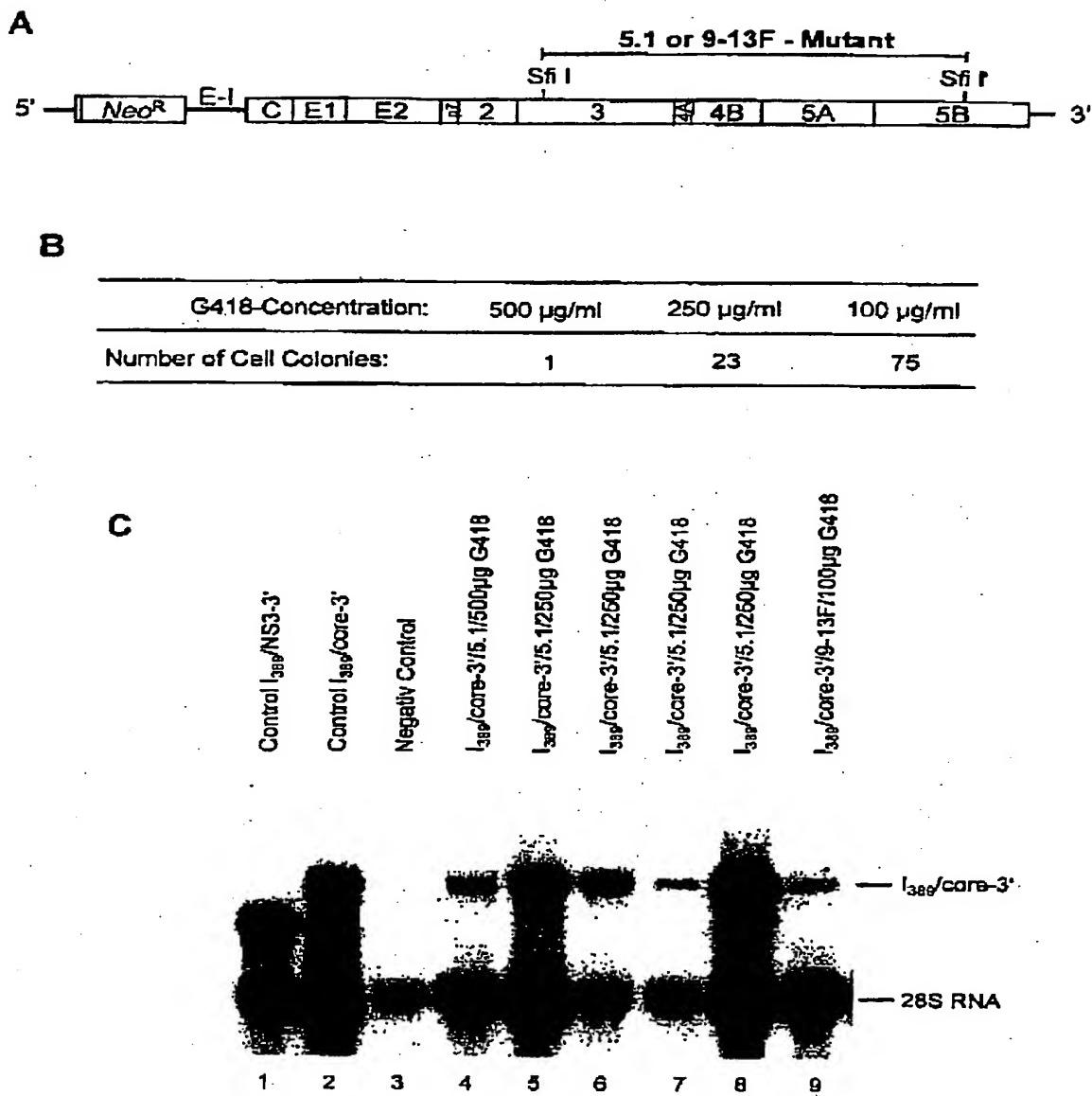


Fig. 10

**A****B****Fig. 11**

**Fig. 12**

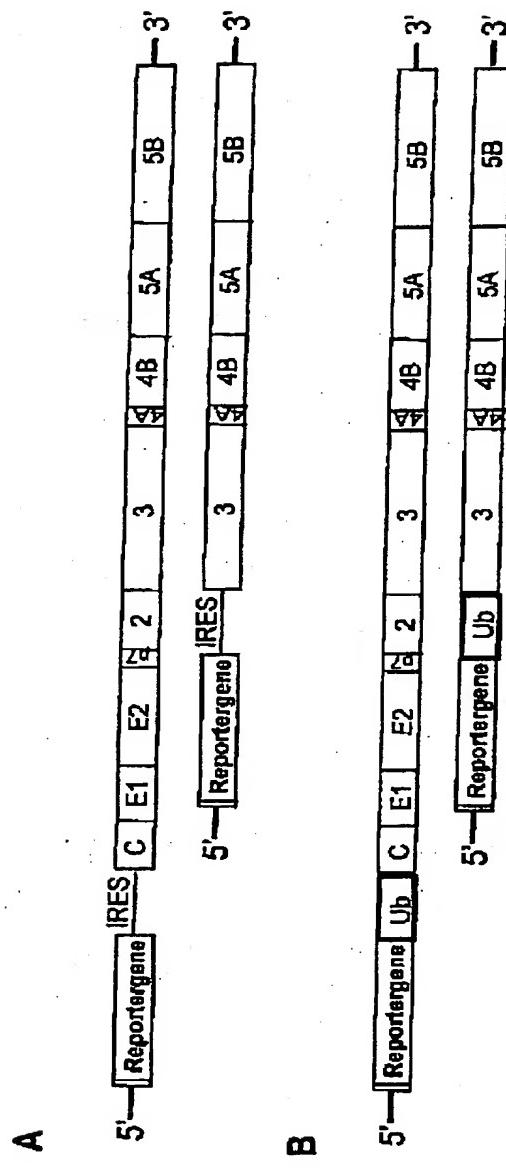


Fig. 13

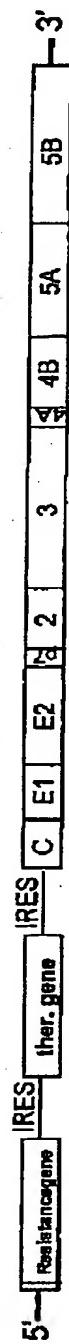
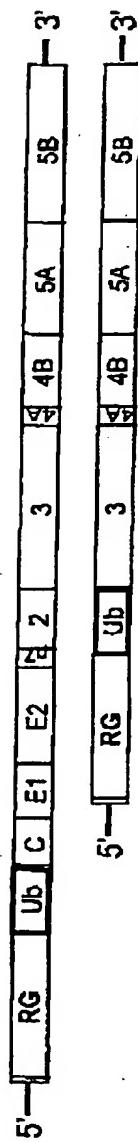


Fig. 14

Fig. 15



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**